

CONVEGNO DI STUDIO
IN MEMORIA DEL PROFESSOR

GIUSEPPE GABRIELLI

Atti della Giornata tenutasi presso il Politecnico di Torino
8 aprile 1991



SUPPLEMENTO 1 AL VOL. 127 - 1993
Atti della Accademia delle Scienze di Torino
Classe di Scienze Fisiche, Matematiche e Naturali

TORINO
ACCADEMIA DELLE SCIENZE
1993



CONVEGNO DI STUDIO
IN MEMORIA DEL PROFESSOR

GIUSEPPE GABRIELLI

Atti della Giornata tenutasi presso il Politecnico di Torino
8 aprile 1991



SUPPLEMENTO 1 AL VOL. 127 - 1993
Atti della Accademia delle Scienze di Torino
Classe di Scienze Fisiche, Matematiche e Naturali

TORINO
ACCADEMIA DELLE SCIENZE
1993

Progetto scientifico e organizzazione

POLITECNICO DI TORINO

ACCADEMIA DELLE SCIENZE DI TORINO

INDICE

	<i>Pagine</i>
Ettore ANTONA, <i>Presentazione</i>	5
Carlo CERCIGNANI, <i>Dinamica dei gas rarefatti: concetti, problemi e metodi di calcolo</i>	7-16
Franco ALGOSTINO, Placido CICALA, <i>Tensori di tensione nelle grandi deformazioni</i>	17-34
Piero PELAGALLI, <i>Ottimizzazione della configurazione di un moderno velivolo a breve raggio</i>	35-79
Attilio SALVETTI, Giovanni MENGALI, <i>Sulla dinamica del volo controllato</i>	81-117
Angelo MIELE, <i>Aircraft Survival in Microburst Encounters</i>	119-165
Ettore ANTONA, <i>La sicurezza nel progetto degli aeromobili e dei veicoli spaziali</i>	167-205
Vittorio GIAVOTTO, Carlo POGGI, Marios CHRYSSANTHOPOULOS, <i>L'instabilità di gusci in materiale composito soggetti a compressione e torsione</i>	207-233
Claudio CANCELLI, <i>Fluidodinamica della cochlea</i>	235-256



PRESENTAZIONE

L'Accademia delle Scienze e il Politecnico di Torino hanno voluto onorare la memoria di Giuseppe Gabrielli, dedicandogli — l'8 aprile 1991 — un «Convegno di studio», la cui organizzazione è stata affidata a un comitato, composto dai professori Rodolfo Zich, Carlo Ferrari, Placido Cicala, Pietro Buzano ed Ettore Antona.

I lavori, svoltisi nelle sede del Politecnico, presenti i familiari e una folta rappresentanza di componenti del mondo Accademico e di quello Industriale, sono stati aperti con parole di saluto del Rettore del Politecnico di Torino e del Presidente dell'Accademia.

*Sono seguiti ricordi della figura e della umanità del Professore Gabrielli, presentati dal Prof. Carlo Ferrari e dagli Ingegneri Fausto Cere-
ti, Franco Giura, Ettore Delmastro e Lucio La Rocca.*

Sono quindi seguite una sessione mattutina e una sessione pomeridiana di lavori scientifici, nelle quali sono state presentate memorie, aventi lo scopo di illustrare lo stato di avanzamento delle conoscenze teoriche e sperimentali, in una ampio insieme di materie connesse con l'ingegneria aeronautica e spaziale.

Nella sessione mattutina sono state svolte quattro relazioni (Cercignani, Algostino-Cicala, Pelagalli, Salvetti) riguardanti la dinamica dei gas, le strutture, l'ottimizzazione nel progetto e la dinamica del volo.

Nella sessione pomeridiana sono state svolte quattro relazioni (Miele, Antona, Giavotto, Cancelli) riguardanti la sopravvivenza in «wind-shear», la sicurezza nel progetto, l'instabilità nelle strutture e la fluidodinamica.

*L'Accademia delle Scienze di Torino si è assunta il compito di curare la stampa delle relazioni, dedicandovi un fascicolo degli Atti, che ricor-
di quella giornata.*

Per il Comitato Organizzatore
Ettore Antona



Dinamica dei gas rarefatti: concetti, problemi e metodi di calcolo

Carlo CERCIGNANI*

Riassunto. Si discutono alcuni problemi tipici che nascono nell'ambito della dinamica dei gas rarefatti, con particolare riguardo per i regimi di rarefazione, per il comportamento della funzione di distribuzione in prossimità dei contorni, ed anche per i metodi di calcolo (di natura sia numerica sia analitica) e i modelli matematici che sono stati introdotti per affrontare tali problemi.

1. Introduzione

Quando si studia il volo nell'alta atmosfera, risulta naturale mettere in forse la validità delle equazioni di Eulero e Navier-Stokes. La teoria cinetica dei gas insegna che tutti i campi di interesse nella gasdinamica si possono ottenere per integrazione d'una funzione di distribuzione $f=f(\mathbf{x}, \boldsymbol{\xi}, \text{tempo } t)$. Così la densità ϱ , la velocità $\mathbf{v}$, le componenti di sforzo p_{ij} (inclusa la pressione p) la corrente termica $\mathbf{q}$, si ottengono (nel caso di un gas monoatomico non troppo denso, a cui ci limiteremo, per semplicità di esposizione) con le formule seguenti [1, 2]:

$$\begin{aligned}\varrho(\mathbf{x}, t) &= \int f(\mathbf{x}, \boldsymbol{\xi}, t) d\boldsymbol{\xi}; \\ \varrho \mathbf{v}(\mathbf{x}, t) &= \int \boldsymbol{\xi} f(\mathbf{x}, \boldsymbol{\xi}, t) d\boldsymbol{\xi}; \\ p_{ij}(\mathbf{x}, t) &= \int c_i c_j f(\mathbf{x}, \boldsymbol{\xi}, t) d\boldsymbol{\xi} \quad (c_i = \xi_i - v_i) \\ p = \varrho RT(\mathbf{x}, t) &= \frac{1}{3} \int c^2 f(\mathbf{x}, \boldsymbol{\xi}, t) d\boldsymbol{\xi} \\ (1.1) \quad \mathbf{q}(\mathbf{x}, t) &= \frac{1}{2} \int c c^2 f(\mathbf{x}, \boldsymbol{\xi}, t) d\boldsymbol{\xi}\end{aligned}$$

* Dipartimento di Matematica, Politecnico di Milano - Milano (Italy).

dove gli indici variano da 1 a 3 e R è la costante dei gas ($=k/m$, essendo m la massa molecolare e k la costante universale di Boltzmann).

Utilizzando solo questa informazione è possibile far vedere che le equazioni di Navier Stokes cessano di essere valide su scale sufficientemente piccole. Si consideri infatti la seguente relazione, conseguenza elementare della formula per p_{ij} data nella (1.1):

$$(1.2) \quad |p_{ij}| \leq \frac{p_{ii} + p_{jj}}{2} \leq 3p/2 .$$

Se fosse valida sempre la relazione costitutiva di Navier Stokes, avremmo per esempio in una corrente con sforzi di taglio dovuti a un gradiente di velocità:

$$(1.3) \quad p_{12} = -\mu \frac{\partial u}{\partial y}$$

($u = v_1$ = componente lungo l'asse x della velocità, μ viscosità) e quindi:

$$(1.4) \quad \left| \mu \frac{\partial u}{\partial y} \right| \leq 3p/2 .$$

Per illustrare il significato della (1.4) notiamo che il libero cammino medio λ è legato a μ dalla relazione:

$$(1.5) \quad \lambda = \frac{\mu}{p} \sqrt{2RT} .$$

e quindi la (1.4) diventa:

$$(1.6) \quad \left| \frac{\partial u}{\partial y} \right| \leq \frac{3}{\lambda} \sqrt{2RT} .$$

Osserviamo che λ è circa 1 metro a un'altezza le circa 130 km. Quindi per gas rarefatti o velocità elevate, le equazioni di Navier-Stokes non sono più compatibili con la teoria cinetica dei gas e si deve usare l'intero apparato di quest'ultima per affrontare la corrispondente dinamica.

2. L'equazione di Boltzmann

L'equazione fondamentale della teoria cinetica dei gas è l'equazione di Boltzmann [1, 2]:

$$(2.1) \quad \frac{\partial f}{\partial t} + \xi \cdot \frac{\partial f}{\partial \mathbf{x}} + \mathbf{X} \cdot \frac{\partial f}{\partial \xi} = Q(f, f)$$

dove il termine d'urto $Q(f, f)$ è dato da:

$$(2.2) \quad Q(f, f) = \frac{1}{m} \iint (f' f'_* - f f_*) B(\theta, |\xi - \xi_*|) d\xi_* d\theta d\epsilon.$$

Qui $B(\theta, |\xi - \xi_*|)$ è un nucleo che contiene i dettagli dell'interazione molecolare, mentre f', f'_*, f_* sono la stessa cosa di f , con la sola differenza che l'argomento ξ è sostituito da ξ', ξ'_*, ξ_* rispettivamente; ξ_* è una variabile d'integrazione (che fisicamente rappresenta la velocità d'una molecola che urta con quelle di velocità ξ , che stiamo seguendo), ξ' e ξ'_* sono le velocità di due molecole che iniziano un urto che le porterà ad avere le velocità ξ and ξ_* . θ and ϵ individuano la direzione di avvicinamento delle due molecole in questione.

L'equazione di Boltzmann (introdotta da Boltzmann nel 1872, dopo importanti studi pionieristici di Maxwell) è divenuta al giorno d'oggi uno strumento pratico nello studio del volo nell'alta atmosfera. Naturalmente, trattandosi di un'equazione integrodifferenziale, dovrà essere risolta con opportune condizioni iniziali e al contorno.

3. Regimi di rarefazione

I parametri adimensionali caratteristici della dinamica dei gas rarefatti sono il numero di Knudsen Kn e il rapporto di velocità molecolare S , definiti da:

$$(3.1) \quad \begin{aligned} Kn &= \lambda/L \quad (L = \text{lunghezza caratteristica}) \\ S &= u/C \quad (u = \text{velocità caratteristica,} \\ &\quad C = \text{velocità termica} = \sqrt{2RT}) \end{aligned}$$

Kn e S sono legati al numero di Mach e di Reynolds, Ma e Re , da:

$$(3.2) \quad Ma = \sqrt{2/\gamma} S; \quad Kn = \sqrt{2\gamma} Ma/Re.$$

Naturalmente si avranno diverse importanti lunghezze macroscopiche come il raggio di curvatura, lo spessore dello strato limite δ , la lunghezza di un corpo, e altrettanti numeri di Knudsen. Particolarmente importante è il numero $Kn_\delta = \lambda/\delta$; quando Kn_δ è maggiore, diciamo, di 0.01, cominciano a sentirsi gli effetti d'uno strato limite sottile presso le pareti, di spessore $\cong \lambda$ (strati di Knudsen) (regime di slittamento). Il gas può ancora essere descritto dalle equazioni di Navier-Stokes, ma bisogna tener conto del fatto che slitta sulla parete con una certa velocità u_s e anche la temperatura subisce una discontinuità. La velocità di slittamento è data da

$$(3.3) \quad u_s = \zeta \left(\frac{\partial u}{\partial n} \right)_w$$

dove ζ è il coefficiente di slittamento, per cui Maxwell [3] aveva trovato il valore approssimato (nel caso di completa diffusione delle molecole alla parete) $\zeta \cong (\sqrt{\pi}/2) \lambda$. In effetti, ζ è circa 15% più grande, come mostrano calcoli più recenti.

Quando il numero di Knudsen cresce avvicinandosi a valori dell'ordine di 1, si assiste a un ispessimento dell'onda d'urto di prora (che ha lo spessore dell'ordine di 6λ) fino ad avere una fusione dello strato d'urto con lo strato limite; si ha allora il cosiddetto regime di strato fuso (in inglese, *merged layer*) o di transizione (tra un comportamento alla Navier-Stokes e uno molecolare libero o quasi libero).

Un ulteriore crescita del numero di Knudsen porta in primo piano l'importanza delle interazioni tra il gas e la superficie solida rispetto ai (pochi urti) tra molecole del gas. Si hanno allora i regimi, già ricordati di molecole libere o quasi-libere. Tutti i regimi di cui si è fatto menzione giocano un ruolo nel volo ad altissima quota (in particolare, nel rientro d'una navetta).

L'interazione tra gas e superfici solide (a cui si può riportare l'origine della resistenza, della portanza e dello scambio termico) è un soggetto di ricerca tipicamente interdisciplinare. Data la difficoltà di modelli fisico-chimici realistici atti a descrivere i fenomeni di adsorbimento e successiva riemissione di molecole gassose da parte d'una parete solida, si usano di solito per descrivere questa interazione modelli semifenome-

nologici, che rispettano i principi generali, ma semplificano drasticamente i dettagli. Il più semplice fu proposto da Maxwell [3] nel 1879 ed è basato sull'idea che una frazione di molecole sia riemessa con riflessione speculare, mentre la parte restante è riemessa in maniera diffusa in tutte le direzioni con distribuzione maxwelliana. Altri modelli usati comunemente sono dovuti a Nocilla [4] e a Cercignani e Lampis [5].

4. Risoluzione dell'equazione di Boltzmann

Come si risolve l'equazione di Boltzmann, già complicata, con condizioni al contorno anch'esse complicate? I metodi analitici si sono rivolti a «correnti semplici» (come quelle di Couette e di Poiseuille) e sono basati su metodi perturbativi, che portano alla cosiddetta equazione di Boltzmann linearizzata [1, 2]. Questi metodi hanno prodotto risultati in accordo spettacolare coi dati sperimentali, ma hanno il difetto di essere applicabili solo a numeri di Mach decisamente subsonici. Tuttavia hanno avuto un ruolo importante nella dinamica dei gas rarefatti; da un lato hanno dato confidenza nella possibilità di usare la descrizione cinetica in problemi di correnti in qualsiasi regime e dall'altro hanno permesso lo studio dei coefficienti di slittamento e di discontinuità termica (per cui la linearizzazione è applicabile anche a numeri di Mach più elevati).

Quando si tien conto della non linearità dell'equazione di Boltzmann, il problema più semplice che si incontra è quello della struttura dell'onda d'urto normale (che, come si è ricordato, è uno strato sottile con spessore $\cong \lambda$, che possiamo supporre perpendicolare all'asse delle x).

A tal fine si può immaginare il problema ambientato nell'intero spazio senza contorni; allora trovare la struttura dell'onda d'urto significa risolvere l'equazione di Boltzmann con le condizioni $s \rightarrow \pm \infty$

$$(4.1) \quad f_0^\pm = q^\pm (2\pi RT^\pm \exp [-(\xi - u^\pm)^2 / (2RT_\pm)]) .$$

È notevole che Mott-Smith [6] sia riuscito a trovare una soluzione approssimata con l'ipotesi rudimentale d'una distribuzione bimodale, ottenuta combinando linearmente le due maxwelliane che descrivono gli stati a monte e a valle:

$$(4.2) \quad f = \nu(x) f_0^+ + (1 - \nu(x)) f_0^- .$$

Questa espressione è compatibile coi bilanci di massa, quantità di moto

ed energia se $\varrho \pm m$, $u \pm$, $T \pm$ soddisfano condizioni di compatibilità che coincidono con quelle ben note di Rankine-Hugoniot.

Il punto critico del metodo è la determinazione del peso $\nu(x)$; questo problema non è stato risolto in maniera soddisfacente, dato il grado di arbitrarietà presente. È notevole però che col metodo comunemente accettato i risultati per le grandezze fisiche, incluso lo spessore, siano più accurati di quelli ottenibili con le equazioni di Navier-Stokes con l'eccezione dei valori del numero di Mach a monte più piccoli (prossimi all'unità).

Un metodo per affrontare i problemi, che ha avuto un certo successo quando non erano disponibili i metodi di simulazione e i calcolatori attuali, è basato sull'idea di sostituire il termine d'urto con qualcosa di più semplice che ne conservi le proprietà (modello cinetico). Il più noto di questi modelli noto con le iniziali dei nomi degli autori che l'hanno proposto, è il modello BGK (Bhatnagar, Gross and Krook [7]), in cui il termine d'urto $Q(f, f)$ è sostituito da

$$(4.3) \quad J(f) = \nu(\Phi - f)$$

dove ν (la frequenza degli urti) è indipendente da ξ ma è proporzionale a ϱ e può dipendere da T , mentre Φ è la maxwelliana locale:

$$(4.4) \quad \Phi = \varrho (2\pi RT)^{-3/2} \exp [- (\xi - v)^2 / (2RT)]$$

ϱ , v , T non sono dati *a priori* ma sono legati a f attraverso le (1.1). I modelli cinetici sono assai utili per sviluppare soluzioni approssimate e idee qualitative, ma non per ottenere risposte dettagliate e precise alle domande poste dagli ingegneri spaziali.

Sono stati quindi sviluppati procedimenti numerici per risolvere l'equazione di Boltzmann coi metodi di discretizzazione tipici dell'analisi numerica e anche metodi per simulare la situazione fisica (metodi Monte Carlo). Questi ultimi sono noti sotto la sigla DSMC (*Direct Simulation Monte Carlo*) che li distingue dalle tecniche Monte Carlo per la valutazione degli integrali d'urto in metodi altrimenti più tradizionali. Ci sono poche limitazioni alla complessità del modello fisico che si può affrontare in questo modo (gas poliatomici, miscele, reazioni chimiche).

In questi metodi gli urti intermolecolari sono trattati su base probabilistica. Il gas viene rappresentato da qualche migliaio di molecole su un calcolatore. Le coordinate e le velocità di queste molecole sono immagazzinate nella memoria e modificate col passar del tempo. I calcoli sono non stazionari. La regione di moto viene suddivisa in piccole celle.

La variabile temporale viene fatta avanzare a passi discreti Δt , piccoli rispetto al libero tempo medio. Questo permette di separare il moto inerziale (o sotto assegnate forze di volume) dal processo d'urto. Il tasso con cui avvengono gli urti è dato da

$$(4.5) \quad r = \sum_{i,j < i} r_{ij}$$

dove il tasso per ogni data coppia (i, j) di molecole è indicato dalla (2.2) ed è quindi dato da:

$$(4.6) \quad r_{ij} = \frac{q}{Nm} \int_{\mathcal{B}} B(\theta, |\xi_i - \xi_j|) d\theta d\epsilon$$

dove N è il numero di particelle usate nella simulazione ξ_i la velocità dell' i -esima molecola ($i = 1, \dots, N$). Ad ogni urto si modificano le velocità (e quindi anche r) col procedimento che ora descriveremo. Sia T_k il tempo tra il $(k-1)$ -esimo e il k -esimo urto. In $[0, \Delta t]$ (per definizione $t=0$ è l'istante dello 0-esimo urto). T_k è scelto così: dopo il $(k-1)$ -esimo urto si «estrae» una coppia (i, j) con distribuzione di probabilità $p_{ij} = r_{ij}/r$ (mantenendo fissa la velocità per ciascuna coppia) e si prende $T_k = 2N/[(N-1)r_{ij}]$. Si «estrae» poi la direzione $\mathbf{k}$ della velocità relativa dopo l'urto $\xi_i^+ - \xi_j^+$ con distribuzione di probabilità $p(\mathbf{k}) = r_{ij}/r$ (i, j e $|\xi_i^+ - \xi_j^+| = |\xi_i - \xi_j|$ fissati). Si sostituiscono poi ξ_i e ξ_j con

$$(4.7) \quad \xi_i^+ = \frac{1}{2} (\xi_i + \xi_j + \mathbf{k} |\xi_i - \xi_j|)$$

$$(4.8) \quad \xi_j^+ = \frac{1}{2} (\xi_i + \xi_j + \mathbf{k} |\xi_i - \xi_j|)$$

Si ripete poi l'operazione $\sum_k T_k$ non supera ΔT .

Sono state proposte molte variazioni del metodo proposto originariamente da Bird [8] (Koura [9], Belotserkovskii e Yanitskii [10], Deshpande [11]), che differiscono solo nel modo di scegliere i T_k . Piuttosto diverso è il metodo di Nanbu [12], che non suddivide Δt e lavora con la probabilità P_i che la i -esima molecola urti nell'intervallo di tempo $[0, \Delta t]$

$$(4.9) \quad P_i = \Delta t \sum_{j=1}^N r_{ij}.$$

Si comincia con r_{ij} valutata per $t=0$ e si estrae un numero a caso ω nell'intervallo $[0, 1]$. A seconda che sia $\omega < P_i$ o $\omega > P_i$, la i -esima molecola urta oppure no nell'intervallo di tempo $[0, \Delta t]$. Se avviene l'urto, si estrae la molecola con cui urta la i -esima molecola con la distribuzione di probabilità:

$$(4.10) \quad P_j^{(i)} = r_{ij} / \sum_{k=1}^N r_{ik}.$$

Si «estrae» poi la direzione $\mathbf{k}$ della velocità relativa dopo l'urto $\xi_i^+ - \xi_j^+$ con distribuzione di probabilità $p(\mathbf{k}) = r_{ij}/r$ (i, j e $|\xi_i^+ - \xi_j^+| = |\xi_i - \xi_j|$ fissati). Si sostituiscono poi ξ_i con ξ_i^* [data dalla prima delle (4.7)]. Si ripete poi il procedimento per tutti gli i .

Il metodo di Nanbu è stato criticato da Koura [13] per una supposta violazione delle leggi fisiche, ma la critica è sbagliata [14]. Quello che è vero è che il metodo non tenta di simulare la dinamica a N corpi, ma piuttosto l'equazione di Boltzmann. Si possono dare dimostrazioni che questa «lotteria» fornisce distribuzioni che convergono in senso opportuno a soluzioni dell'equazione di Boltzmann. Tipico è il seguente:

Teorema [15, 16]: Se l'equazione di Boltzmann con dato iniziale f_0 ha una soluzione liscia, non negativa $f(\mathbf{x}, \xi)$ integrabile, allora la soluzione $\hat{f}$ del metodo di Nanbu converge debolmente nello spazio delle funzioni integrabili a f , nel senso che, per ogni funzione di prova ϕ limitata,

$$(4.11) \quad \phi(\mathbf{x}, \xi) \int \phi \hat{f} d\mathbf{x} d\xi \rightarrow \int \phi f d\mathbf{x} d\xi \quad \text{quando } N \rightarrow \infty, \Delta x, \Delta t \rightarrow 0.$$

L'entità del lavoro di calcolo è proporzionale a N per il metodo di Bird, a N^2 per quello di Nanbu. Però Babovsky ha indicato come ridurre questo lavoro per il metodo di Nanbu suddividendo l'intervallo $[0, 1]$ in N sottointervalli eguali. Non solo teoricamente ma anche praticamente il lavoro di calcolo risulta confrontabile con quello di Bird [17].

5. Turbolenza e dinamica dei gas rarefatti

In quest'ultimo paragrafo vorrei esaminare la questione se ci sia un terreno significativo in cui la turbolenza di un gas debba essere descritta con l'equazione di Boltzmann.

Un'asserzione comune è che le scale della turbolenza e quelle dei moti molecolari siano ben separate. Questo può essere sempre vero per un liquido, ma richiede un esame più accurato nel caso d'un gas. Si consideri infatti la scala dissipativa della turbolenza ℓ_D e il libero cammino medio λ ; si ha:

$$(5.1) \quad \frac{\lambda}{\ell_D} = \frac{\lambda}{L} (\text{Re})^{3/4} = \text{Kn} (\text{Re})^{3/4} = (\text{Ma})^{3/4} (\text{Kn})^{1/4}$$

dove L è una scala macroscopica e si è tenuto conto della relazione $\text{Re} = \text{Ma}/\text{Kn}$. Se prendiamo $\text{Re} = 10^4$ per la turbolenza pienamente sviluppata, otteniamo che λ e ℓ_D sono dello stesso ordine per $\text{Kn} = 10^{-3}$ e $\text{Ma} = 10$. Così nelle correnti ipersoniche moderatamente rarefatte le scale di turbolenza e il libero cammino medio sono dello stesso ordine di grandezza!

Queste stime preliminari sono da prendere con la dovuta cautela (a causa della possibile presenza di fattori «dell'ordine dell'unità» non proprio vicini ad uno). Ma queste considerazioni devono essere tenute presenti nello studio di correnti ad alti numeri di Mach. Si tratta per altro d'un settore completamente inesplorato, che dovrebbe essere oggetto di esame sia teorico che numerico.

BIBLIOGRAFIA

- [1] CERCIGNANI C., *The Boltzmann equation and its applications*, Springer, New York, 1988.
- [2] CERCIGNANI, C., *Mathematical Methods in Kinetic Theory*, Plenum Press, revised edition, New York, 1990.
- [3] MAXWELL J.C., *Phil. Trans. Royal Soc. I*, Appendix (1879).
- [4] NOCILLA S. in *Rarefied Gas Dynamics*, L. Talbot, ed., 169, Academic Press, New York (1961).
- [5] CERCIGNANI C., LAMPIS M., *Transport Theory and Stat. Phys.* **1**, 101 (1971).
- [6] MOTT-SMITH H.M., *Phys. Rev.* **82**, 885, 1951.
- [7] BHATNAGAR P.L., GROSS E.P., KROOK M., *Phys. Rev.* **94**, 511 (1954).
- [8] BIRD G.A., *Phys. Fluids* **13**, 2676, 1970.
- [9] KOURA K., *Phys. Fluids* **13**, 1457, 1970.
- [10] BELOTSEKOVSKII O.M., YANITSKII V., *Zhurn. Vych. Mat. Mat. Fiz.* **15**, 1195, 1975 (in russo).
- [11] DESHPANDE S.M., Dept. Aero. Engng. Indian Inst. Science Rep. **78**, FM4, 1978.
- [12] NANBU K., *J. Phys. Soc. Japn.* **49**, 2042, 1980.
- [13] KOURA K., *J. Phys. Soc. Japn.* **50**, 3829, 1981.
- [14] NANBU K., *J. Phys. Soc. Japn.* **50**, 3831, 1981.
- [15] BABOVSKY H., *European Journal of Mechanics B/Fluids* **8**, 41, 1989.
- [16] BABOVSKY H., ILLNER R., *SIAM Journal of numerical analysis* **26**, 45, 1989.
- [17] GROPENIEßER F., NEUNZERT H., STRUCKMEIER J., in *Applied and Industrial Mathematics*, R. Spigler, ed., p. 111, Kluwer, Dordrecht, 1991.

Tensori di tensione nelle grandi deformazioni

Franco ALGOSTINO*, Placido CICALA**

Summary. Stress tensors are presented in dyadic form in connection with strain tensors for large deformations. Their properties related to the constitutive relationships for generic materials are examined.

1. Introduzione

L'attuale sviluppo degli strumenti di calcolo numerico ha reso possibile la risoluzione di problemi lineari e nonlineari della Meccanica del continuo, di complessità prima proibitiva. Sorge perciò per l'ingegnere la necessità di approfondire le conoscenze della materia ad evitare che egli sia mero esecutore di programmi costruiti da pochi esperti. Esiste nel campo un cospicuo trattato [1], peraltro difficilmente utilizzabile per la vastità e il livello della trattazione. Su quello poggiano testi più recenti che in parte coprono l'esigenza accennata [2].

Nell'analisi strutturale, come in tutte le questioni di Fisica matematica, il calcolo tensoriale mostra le sue peculiari capacità di sintesi, specialmente evidenti nell'impiego di coordinate curvilinee, in molti casi indispensabili. In lavori precedenti [3, 4] si è notato che si conseguono notevoli vantaggi di chiarezza ponendo i tensori in forma diadica, ossia scrivendo le componenti insieme alle terne di vettori costituenti le basi a cui le componenti vanno riferite: altrimenti possono sorgere ambiguità quando siano da usare più riferimenti. Di fatti in un'importante memoria sulla risoluzione di problemi di grandi deformazioni [5] sono state segnalate imprecisioni in diversi lavori del settore. Con la scrittura esplicita si opera sempre su vettori pur sfruttando i vantaggi del calcolo tensoriale. Grazie alle convenzioni sugli indici, le espressioni conservano in buona misura sinteticità¹.

* Dipartimento di Ingegneria del Politecnico di Torino.

** Emerito di Scienza delle Costruzioni nel Politecnico di Torino.

¹ Si scrivono in caratteri greci gli indici di co- e controvarianza: ad essi si applicano le convenzioni per indici ripetuti. Invece gli indici a, b, c stanno per le permutazioni 1, 2, 3 e 2, 3, 1 e 3, 1, 2 senza sommatorie anche se ripetuti: quando deve attuarsi la sommatoria alle tre scelte si applica il simbolo Σ . Tutti gli altri indici in caratteri latini sono distintivi non numerali, mentre l'indice τ segnala la trasposizione dei tensori. I tensori sono di 2° ordine (diadiche).

Seguendo questo indirizzo, qui si esaminano vari tensori di tensione e si studia in forma generale la correlazione con i tensori di deformazione come può essere stabilita dall'esperienza o dalle ipotesi sostitutive.

2. Proprietà del campo di tensioni

Nel punto generico P del corpo, $\mathbf{x}$ è il vettore di posizione, funzione delle variabili x_h^{α} ¹⁾, e $\mathbf{T}$ il tensore di tensione. Le proprietà dello stato di tensione, nella notazione tensoriale, sono compendiate nelle relazioni

$$(1) \quad \mathbf{T} \cdot \mathbf{n} = \boldsymbol{\sigma}, \quad \text{div } \mathbf{T} + \mathbf{f}_v = 0, \quad \mathbf{T} = \mathbf{T}_\tau.$$

La $(1)_1$ dà il vettore tensione risultante riferito all'area della sezione dalla quale esce normalmente il versore $\mathbf{n}$. La $(1)_2$ è la condizione di equilibrio delle forze sull'elemento, ossia la esterna, prodotto del volume per $\mathbf{f}_v$, e la risultante delle tensioni sul contorno dell'elemento. La $(1)_3$ impone la condizione di equilibrio alle rotazioni dell'elemento eguagliando il tensore al suo trasposto.

Per specificare le (1) in forma vettoriale si calcola in P la terna

$$(2) \quad \mathbf{h}_\alpha = \partial \mathbf{x} / \partial x_h^\alpha$$

e da questa la reciproca e il prodotto misto

$$(3) \quad \mathbf{h}^a = \mathbf{h}_b \times \mathbf{h}_c / j_h, \quad j_h = \mathbf{h}_1 \cdot \mathbf{h}_2 \times \mathbf{h}_3.$$

Con un vertice in P si costruisce il parallelepipedo di lati $\mathbf{h}_a dx^a$ (fig. 1); si indica con $s^a dx_h^b dx_h^c$ la risultante delle tensioni sulla faccia attraversata dal vettore $\mathbf{h}_a$ uscente. Con quelle risultanti si costruisce la diadica

$$(4) \quad \mathbf{S} = s_h^\alpha \mathbf{h}_\alpha = j_h \mathbf{T}.$$

Così le (1) sono poste in forma vettoriale

$$(5) \quad \boldsymbol{\sigma} = s_h^\alpha \mathbf{h}_\alpha \cdot \mathbf{n} / j_h, \quad \partial s_h^\alpha / \partial x_h^\alpha + j_h \mathbf{f}_v = 0, \quad s_h^\alpha \times \mathbf{h}_\alpha = 0.$$

¹⁾ L'apice di controvarianza, propriamente, non spetta alle coordinate generali, ma ai loro incrementi.

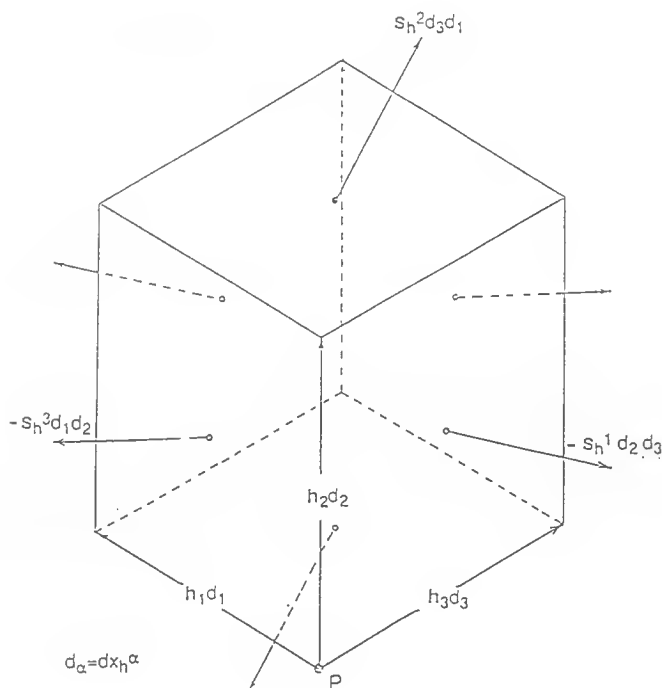


Figura 1

A giustificare la (5)₁ si tracciano i 3 vettori $\mathbf{h}_a d_a$ che portano da P ai punti Q_a . L'area triangolare $Q_1 Q_2 Q_3$ viene rappresentata mediante il vettore perpendicolare

$$(6) \quad \mathbf{n}A = j_h \Sigma \mathbf{h}^a d_b d_c / 2$$

ottenuto dal semiprodotto vettoriale di due lati del triangolo. L'equilibrio delle forze sul tetraedro $PQ_1Q_2Q_3$ dà

$$(7) \quad \sigma A = \Sigma s_h^\alpha d_b d_c / 2 .$$

Sostituendo nella (7) il valore $d_b d_c$ fornito dalla (6) moltiplicata per $\mathbf{h}_a$ si ottiene la (5)₁.

Alla (5)₂ si arriva considerando che le risultanti delle tensioni sulle due facce del parallelepipedo attraversate da $\mathbf{h}_a$ differiscono per la forza $dx_h^b dx_h^c (\partial s_h^a / \partial x_h^a) dx_h^a$. Il prodotto dei tre incrementi di coordina-

te e di j_h dà il volume del parallelepipedo: questo, moltiplicato per $\mathbf{f}_v$, dà il carico. Quindi i due addendi nella (5)₂, moltiplicati per gli incrementi, rappresentano le forze in equilibrio. La stessa equazione si ricava dalla (1)₂ considerando che è $\text{div } \mathbf{T} = \partial \mathbf{T} / \partial x_h^\alpha \cdot \mathbf{h}^\alpha$ ed è

$$(8) \quad \partial j_h / \partial x_h^\alpha = j_h \partial \mathbf{h}_\alpha / \partial x_h^\beta \cdot \mathbf{h}^\beta .$$

L'equazione (5)₃ si giustifica considerando che le risultanti delle tensioni su facce opposte, nei loro termini fondamentali, costituiscono forze uguali opposte, applicate agli estremi di un segmento uguale ai lati non contenuti nelle due facce.

Le accennate deduzioni delle (5) si allacciano a concetti propri dei corsi di base: peraltro esse vanno completate da considerazioni sugli ordini di infinitesimo dei vari contributi, per constatare che tutti i termini fondamentali sono tenuti in conto.

Le risultanti vengono scisse in componenti controvarianti scrivendo

$$(9) \quad \mathbf{s}_h^\alpha = s_h^{\alpha\beta} \mathbf{h}'_\beta = j_h t_h^{\alpha\beta} \mathbf{h}'_\beta .$$

Si introduce qui una terna $\mathbf{h}'$, diciamo di scomposizione, per distinguerla dalla $\mathbf{h}$ di «sezionamento» che stabilisce le giaciture a cui le risultanti si riferiscono. Nella (9), come è generalmente in uso, il primo apice della componente individua il piano di sezione, il secondo la direzione di scomposizione. Invece, esplicitando il tensore con le due basi, queste appaiono in ordine inverso, come vuole la (5)₁:

$$(10) \quad \mathbf{T} = \mathbf{h}'_\beta t_h^{\alpha\beta} \mathbf{h}_\alpha .$$

Si noti che solo la terna di sezionamento è legata alle variabili indipendenti secondo la (2). Se si adotta la stessa terna per le due basi, risulta dalla (1)₃ $t_h^{\alpha\beta} = t_h^{\beta\alpha}$. Perciò in questo caso gli indici delle basi si scrivono nello stesso ordine di quelli delle componenti.

3. Forme equivalenti del tensore di tensione

Nel corpo conviene stabilire, in uno stato di riferimento, che conveniamo di dire «iniziale», un sistema di coordinate ξ^α , di regola correlato alla forma del corpo: queste coordinate vengono legate ai punti materiali e si ritrovano nella configurazione attuale (deformata) costituendo un riferimento naturale o «materiale» atto alla rappresentazione

lagrangiana della deformazione. Nello spazio si dispone inoltre un sistema di coordinate x^α e, talvolta in aggiunta, uno di coordinate cartesiane ortogonali y^α .

Si indicano con $\mathbf{g}_\alpha$, $\mathbf{e}_\alpha$, $\mathbf{i}_\alpha$ le basi che si ottengono dalla (2) sostituendo alle x_h^α rispettivamente le coordinate ξ^α , x^α , y^α : dalle (3) si hanno le reciproche $\mathbf{g}^\alpha$, $\mathbf{e}^\alpha$, $\mathbf{i}^\alpha = \mathbf{i}_\alpha$ (quest'ultima ortonormale) e i prodotti misti $j, j_e, 1$. In quanto i tre riferimenti rappresentano la medesima configurazione attuale, uno dei tre sistemi di coordinate può definire gli altri due. In particolare, dalle derivazioni $\partial \mathbf{x}(x^\beta)/\partial \xi^\alpha$, $\partial \mathbf{x}(\xi^\beta)/\partial x^\alpha$ si hanno le relazioni

$$(11) \quad \mathbf{g}_\alpha = \mathbf{e}_\beta x_\alpha^\beta, \quad \mathbf{e}_\alpha = \mathbf{g}_\beta \xi_\alpha^\beta$$

$$(12) \quad x_\alpha^\beta = \partial x^\beta / \partial \xi^\alpha, \quad \xi_\alpha^\beta = \partial \xi^\beta / \partial x^\alpha.$$

Moltiplicando le (11)₁, (11)₂ per $\mathbf{e}^\lambda$, $\mathbf{g}^\lambda$ si ottengono le espressioni

$$(13) \quad x_\alpha^\beta = \mathbf{e}^\beta \cdot \mathbf{g}_\alpha, \quad \xi_\alpha^\beta = \mathbf{g}^\beta \cdot \mathbf{e}_\alpha$$

per gli elementi delle relative matrici di Jacobi. Le stesse si ottengono moltiplicando per $\mathbf{g}_\lambda$, $\mathbf{e}_\lambda$ le relazioni conseguenti

$$(14) \quad \mathbf{e}^\beta = \mathbf{g}^\alpha x_\alpha^\beta, \quad \mathbf{g}^\beta = \mathbf{e}^\alpha \xi_\alpha^\beta.$$

Ai tre riferimenti corrispondono per il tensore di tensione le espressioni equivalenti, con le due basi uguali,

$$(15) \quad \mathbf{T} = \mathbf{g}_\alpha t^{\alpha\beta} \mathbf{g}_\beta = \mathbf{e}_\alpha t_e^{\alpha\beta} \mathbf{e}_\beta = \mathbf{i}_\alpha \sigma^{\alpha\beta} \mathbf{i}_\beta.$$

Effettuando i prodotti $\mathbf{e}^\lambda \cdot \mathbf{T} \cdot \mathbf{e}^\mu$ sulle prime due espressioni e adottando le (13)₁ si ottiene l'espressione

$$(16) \quad t_e^{\lambda\mu} = x_\alpha^\lambda t^{\alpha\beta} x_\beta^\mu.$$

Questa è la relazione, tipica dei tensori detti «assoluti», fra le matrici delle componenti e quelle di Jacobi del cambio di riferimento. Per i tensori «relativi di peso 1» come $\mathbf{S}$ nella (4), si deve passare ai corrispondenti assoluti dividendo per il prodotto misto della terna di sezionamento.

Con qualunque base di sezionamento, può risultare conveniente effettuare la scomposizione (9) su assi cartesiani:

$$(17) \quad \mathbf{s}_h^\alpha = s_n^\alpha{}_\beta \mathbf{i}^\beta.$$

Con ciò si riduce immediatamente a forma scalare la (5)₂:

$$(18) \quad \partial s_n^\alpha / \partial x_h^\alpha + j_h \mathbf{f}_v \cdot \mathbf{i}_\beta = 0.$$

Peraltro, per le 9 componenti s_n^α che sostituiscono le 6 delle formulazioni (15) si debbono introdurre le 3 equazioni scalari che si deducono dalla (5)₃. Le stesse considerazioni valgono anche se la terna $\mathbf{i}^\beta$ è sostituita da una non ortonormale: la sua reciproca entra nella (18) al posto di $\mathbf{i}_\beta$ nell'ultimo termine.

4. Tensori delle deformazioni finite

Per l'esame della trasformazione che l'intorno di P subisce nel passaggio dello stato di riferimento all'attuale vale la diadica

$$(19) \quad \mathbf{F} = \mathbf{g}_\alpha \mathbf{g}^{\alpha*}$$

dove $\mathbf{g}^{\alpha*}$ è la terna reciproca della $\mathbf{g}_\alpha^* = \partial \mathbf{x}^* / \partial \xi^\alpha$, derivata dal vettore di posizione nello stato di riferimento. Il prodotto $\mathbf{g}^{\alpha*} \cdot d\mathbf{x}^*$ rappresenta l'incremento della coordinata ξ^α corrispondente al vettore $d\mathbf{x}^*$: quindi $\mathbf{F} \cdot d\mathbf{x}^* = d\mathbf{x}$ dà l'elemento corrispondente nella deformazione.

Dal tensore $\mathbf{F}$ si ricavano i tensori

$$(20) \quad \mathbf{C} = \mathbf{F}_\tau \cdot \mathbf{F} = \mathbf{g}^{\alpha*} g_{\alpha\beta} \mathbf{g}^{\beta*}$$

$$(21) \quad \mathbf{E} = \mathbf{g}^{\alpha*} \gamma_{\alpha\beta} \mathbf{g}^{\beta*}$$

essendo $g_{\alpha\beta} = \mathbf{g}_\alpha \cdot \mathbf{g}_\beta$, $\gamma_{\alpha\beta} = (g_{\alpha\beta} - g_\alpha^* \cdot g_\beta^*)/2$.

Sia $\mathbf{p}_\alpha^*$ la terna dei versori ortogonali (principali di riferimento) a cui nella deformazione corrisponde la terna $\lambda_a \mathbf{p}_a$ nello stato attuale, essendo $\mathbf{p}_\alpha$ terna ortonormale: $\lambda_a = L_a / L_a^*$ misura il rapporto fra le lunghezze degli elementi corrispondenti. In base a queste determinazioni si dà ai tensori in gioco la forma:

$$(22) \quad \mathbf{F} = \mathbf{R} \cdot \mathbf{U}, \quad \mathbf{R} = \mathbf{p}_\alpha \mathbf{p}_\alpha^*, \quad \mathbf{U} = \sum \mathbf{p}_a^* \lambda_a \mathbf{p}_a^*$$

$$(23) \quad \mathbf{C} = \sum \mathbf{p}_a^* \lambda_a^2 \mathbf{p}_a^*, \quad 2\mathbf{E} = \sum \mathbf{p}_a^* (\lambda_a^2 - 1) \mathbf{p}_a^*.$$

Secondo le relazioni (22) la trasformazione che porta dallo stato iniziale all'attuale si scompone nei passi seguenti:

- a) una deformazione «in loco» (stretching) rappresentata dalla diadica $\mathbf{U}$, in cui le fibre dirette secondo i versori $\mathbf{p}_\alpha^*$ si allungano nel rapporto λ_α se questo è > 1 (o si accorciano, se < 1) senza cambiare direzione;
- b) un moto rigido contenente la rotazione $\mathbf{R}$ in cui i versori $\mathbf{p}_\alpha^*$ si mutano nei $\mathbf{p}_\alpha$.

Ovviamente tale scomposizione è idealmente possibile per il corpo frazionato in particelle: la deformazione $\mathbf{U}$ in genere non verifica la congruenza. Tuttavia, nell'esame delle proprietà del corpo, la scomposizione ha notevole importanza.

Si osserva che nei tensori $\mathbf{C}$ e $\mathbf{E}$ gli effetti della rotazione scompaiono. Corrispondentemente le componenti indipendenti dei tensori simmetrici anzidetti si riducono a 6 dalle 9 che definiscono $\mathbf{F}$.

In molte applicazioni di Ingegneria il moto rigido $\mathbf{x}-\mathbf{x}^*$ associato alle rotazioni $\mathbf{R}$ assume entità assai maggiore che le componenti della deformazione; allora per queste sono lecite le linearizzazioni.

5. L'equazione dei lavori virtuali

Le forze applicate al corpo, per gli spostamenti virtuali $\delta\mathbf{x}$ dalla configurazione attuale, compiono il lavoro «esterno»

$$(24) \quad L_e = \int \mathbf{f}_v \cdot \delta\mathbf{x} \, dV + \int \boldsymbol{\sigma} \cdot \delta\mathbf{x} \, dS.$$

Il primo integrale è esteso al volume del corpo o alla parte di esso che subisce quegli spostamenti senza dislocazioni, il secondo alla relativa superficie. L'equazione in oggetto eguaglia L_e al lavoro «interno» L_i che si calcola scomponendo il corpo in elementi di volume e sommando i contributi derivanti dalle forze, in equilibrio su ogni elemento interno o di contorno. Per il parallelepipedo di lati $\mathbf{h}_a dx_h^a$ le forze $\pm s_h^a dx_h^b dx_h^c$ sulle facce attraversate da $\mathbf{h}_a$ lavorano per lo spostamento relativo $(\partial\delta\mathbf{x}/\partial x_h^a) dx_h^a$: sostituendo al prodotto degli incrementi il rapporto dV/j_h si trova

$$(25) \quad L_i = \int s_h^\alpha \cdot (\partial\delta\mathbf{x}/\partial x_h^\alpha) \, dV/j_h.$$

Si deve notare che la differenza fra le risultanti delle tensioni sulle facce opposte lavora per lo spostamento $\delta\mathbf{x}$ e dà pure un contributo di ordine dV : questo però è compensato dal carico sull'elemento.

L'espressione (25) è generale sia per la scelta delle coordinate sia per la scomposizione delle tensioni. Se si adottano le (17), dal prodotto scalare nella (25) si ottengono 9 addendi $s_h^{\alpha\beta} \partial \delta y^\beta / \partial x_h^\alpha$. Con qualsiasi scomposizione (15), in virtù della simmetria delle componenti $t_h^{\alpha\beta} = s_h^{\alpha\beta} / j_h$, l'integrando della (25) può porsi nella forma:

$$(26) \quad dL_i/dV = t_h^{\alpha\beta} (\mathbf{h}_\alpha \cdot \partial \delta \mathbf{x} / \partial x_h^\beta + \mathbf{h}_\beta \cdot \delta \mathbf{x} / \partial x_h^\alpha) / 2 .$$

Con le coordinate $x_h^\alpha = \xi^\alpha$, posto che queste non siano modificate dagli spostamenti virtuali, risulta $\partial \delta \mathbf{x} / \partial \xi^\alpha = \delta \mathbf{g}_\alpha$ e quindi

$$(27) \quad dL_i/dV = t^{\alpha\beta} \delta \gamma_{\alpha\beta} .$$

Poiché $\delta \mathbf{g}_\alpha^* = 0$, in accordo con la (21) si scrive

$$(28) \quad \delta \mathbf{E} = \mathbf{g}^{\alpha*} \delta \gamma_{\alpha\beta} \mathbf{g}^{\beta*} .$$

Introducendo il tensore gradiente degli spostamenti virtuali

$$(29) \quad \delta \mathbf{F}_o = \partial \delta \mathbf{x} / \partial x_h^\alpha \mathbf{h}^\alpha$$

e il tensore $\delta \mathbf{E}_0$, incremento di deformazione per $\mathbf{g}^{\alpha*} = \mathbf{g}^\alpha$, si ha

$$(30) \quad dL_i/dV = \mathbf{T} : \delta \mathbf{F}_0 = \mathbf{T} : \delta \mathbf{E}_0$$

indicandosi con: il doppio prodotto scalare di due tensori, contenente i prodotti delle prime due terne e delle seconde. Delle due espressioni alternative, la prima riconduce alla (25) attraverso le (4) e (29): la seconda va interpretata mediante la (26) oppure mediante la relazione conseguente $\delta \mathbf{E}_0 = (\delta \mathbf{F}_0 + \delta \mathbf{F}_{0T}) / 2$. Si noti che, sebbene sia $\partial \delta \mathbf{x} / \partial x_h^\alpha \neq \delta \mathbf{h}_\alpha$, i prodotti $\mathbf{h}_\alpha \cdot \delta \mathbf{E}_0 \cdot \mathbf{h}_\beta$ rappresentano le componenti della deformazione virtuale nel riferimento generico x_h^α .

6. Costruzione delle leggi costitutive

La successione di configurazioni attuata dal corpo nel processo di deformazione in esame sia definita assegnando il vettore $\mathbf{x}$ come funzione delle coordinate materiali e di un parametro t , sempre crescente, che denominiamo tempo anche se si considerano fenomeni non influenzati dai reali intervalli fra successivi eventi. A partire dalla configurazione $\mathbf{x}^*$

di riferimento, in ciascun istante il tensore $\mathbf{E}$ rivela il cambiamento di posizione relativa di punti contigui, deformazione dell'intorno di P riferita alle basi $\mathbf{g}_\alpha^*$. Nell'indagine sulla correlazione fra quella deformazione e gli stati conseguenti di tensione di regola si ammette che la correlazione abbia carattere locale: ossia essa non sia influenzata dal gradiente della deformazione $\mathbf{E}$. Allora, limitando lo studio a deformazioni isoterme, per dare alle leggi di correlazione un fondamento fisico si dovrebbe disporre di un campione avente le caratteristiche della particella in esame e sottoporre il provino, a temperatura costante, alla sequenza di deformazioni $\mathbf{E}$ uniformemente distribuite che si susseguono nel punto P . Se al tempo t la terna $\mathbf{g}_\alpha^*$ nel provino si trasforma nella $\mathbf{g}_{s\alpha}$, questa sarà congruente alla $\mathbf{g}_\alpha$ nel corpo, ossia sarà $\mathbf{g}_{\alpha\beta} = \mathbf{g}_{s\alpha} \cdot \mathbf{g}_{s\beta}$ e il tensore di tensione assumerà il valore

$$(31) \quad \mathbf{T}_s = \mathbf{g}_{s\alpha} t^{\alpha\beta} \mathbf{g}_{s\beta}$$

con le stesse componenti del tensore in P , dato da

$$(32) \quad \mathbf{T} = \mathbf{R}_s \cdot \mathbf{T}_s \cdot \mathbf{R}_{sT}$$

essendo $\mathbf{R}_s = \mathbf{g}_\alpha \mathbf{g}_{s\alpha}^\alpha$ la rotazione che porta la terna $\mathbf{g}_{s\alpha}$ a coprire la $\mathbf{g}_\alpha$. Questa rotazione non ha alcuna influenza sul fenomeno.

Per realizzare idealmente il procedimento si assegni a $\mathbf{R}_s$ il valore $(22)_2$: allora risulta $\mathbf{U} = \mathbf{R}_s^{-1} \cdot \mathbf{F} = \mathbf{g}_{s\alpha} \mathbf{g}_\alpha^*$, ossia la deformazione attuata nel modello è la $\mathbf{U}$ che lascia invariata la direzione della terna principale, trasformando la terna $\mathbf{g}_\alpha^*$ nella $\mathbf{g}_{s\alpha}$. Su questa seconda si trovano le componenti del tensore che, mediante un moto rigido porta allo stato attuale di tensione. La deformazione per similitudine $\mathbf{U}$ (stretching) è definita dai tre rapporti λ_α di dilatazione e dai tre che precisano la terna $\mathbf{p}_\alpha^*$ (ad esempio, i tre angoli di Eulero). Dunque la legge costitutiva del materiale è individuata quando si conoscano le 6 componenti di tensione come funzioni dei 6 parametri anzidetti.

Stanti le difficoltà di attuazione delle misure nel campo delle grandi deformazioni, si è in generale preferito postulare modelli di comportamento plausibile per materiali in uso e controllare le conclusioni, quando è possibile. In questo indirizzo ha vasta applicazione la teoria del materiale «iperelastico», capace di accumulare energia nel processo di deformazione finita e di restituirla interamente quando lo si riporti allo stato iniziale. L'iperelasticità comporta una corrispondenza biunivoca fra tensioni e deformazioni, con un bilancio in cui entra l'energia termica. Nel modo più semplice per la teoria e l'esperimento il processo si

attua a temperatura costante, come qui si pone, si da escluderne gli effetti nel bilancio energetico.

Detta ϱ la densità, Φ l'energia elastica riferita alla massa e quindi $\varrho dV \delta \Phi$ l'equivalente del lavoro dL_i , dalla (25) si ha

$$(33) \quad s^\alpha \cdot \delta g_\alpha = j \varrho \delta \Phi .$$

Se si adotta la scomposizione (17) l'energia Φ va espressa in funzione delle quantità $g_\alpha \cdot i^\beta = y^\beta_\alpha$ con la notazione (13)₁ per coordinate cartesiane: per la scomposizione (15)₁ entrano le componenti $\gamma_{\alpha\beta}$ di deformazione. La variazione di dette variabili è indipendente: quindi si desumono 9 oppure 6 equazioni rispettivamente:

$$(34) \quad s_n^{\alpha\beta} = j \varrho \partial \Phi / \partial y^\beta_\alpha, \quad t^{\alpha\beta} = \varrho \partial \Phi / \partial \gamma_{\alpha\beta} .$$

Qualora nella (34)₁ si esprima Φ in funzione delle $g_{\alpha\beta} = y^\lambda_\alpha y^\lambda_\beta$, le condizioni (5)₃ risultano direttamente verificate. Infatti si ha

$$y^a_\mu \partial \Phi / \partial y^b_\mu = (y^a_\alpha y^b_\beta + y^a_\beta y^b_\alpha) \partial \Phi / \partial \gamma_{\alpha\beta} .$$

Poiché l'espressione in parentesi non cambia se si permutano gli apici a, b risulta $s_n^{\mu a} y^b_\mu = s_n^{\mu b} y^a_\mu$ come si ha dalla (5)₃ moltiplicata scalarmente per i_c .

Nel riferimento alla terna p_α , se il corpo, inizialmente scarico, è isotropo, Φ non contiene le componenti $\gamma_{\alpha\beta}$ per $\alpha \neq \beta$: quindi, secondo la (34)₂ le direzioni principali per le tensioni coincidono con quelle della deformazione.

7. Tensori di Hill

È detto «secondo tensore di Piola-Kirchhoff» quello definito dalla diadica relativa

$$(35) \quad S'' = g_\alpha^* s^{\alpha\beta} g_\beta^* .$$

Questa espressione mostra che così le componenti controvarianti, dalla base g_α sono riportate sulla base g_α^* : diciamo allotropico il nuovo tensore che così si crea. Esso presenta la proprietà di dare con l'incremento della deformazione finita il doppio prodotto $S'' : \delta E$ uguale al valore che si ha per $S : \delta E_0$. Secondo quanto suggerito da Hill, si possono costruire

quanti si vogliano tensori «associati» S_e , E_e ponendo

$$(36) \quad E_e = \sum p_a^* e_a p_a^*$$

con e_a funzione assegnata di λ_a e imponendo a S_e l'uguaglianza

$$(37) \quad S_e : \dot{E}_e = S : \dot{E}_o$$

scrivendosi qui $\dot{\psi} = \delta\psi/\delta t$ al posto della variazione $\delta\psi$ anche se arbitraria. Si rappresenta la velocità rotatoria della terna p_α^* con il tensore

$$(38) \quad \dot{R}^* = \dot{p}_\alpha^* p_\alpha^* = p_\alpha^* \omega_{\alpha\beta}^* p_\beta^*$$

antisimmetrico. Dalla espressione (36), derivata come un prodotto, si ha

$$(39) \quad \dot{E}_e = \dot{R}^* \cdot E_e + \sum p_a^* \dot{e}_a p_a^* + E_e \cdot \dot{R}^* .$$

Ponendo $e_a' = de_a/d\lambda_a$ si hanno le componenti del tensore simmetrico E_e

$$(40) \quad p_a^* \cdot \dot{E}_e \cdot p_a^* = e_a' \dot{\lambda}_a, \quad p_a^* \cdot \dot{E}_e \cdot p_b^* = (e_b - e_a) \omega_{ab}^* .$$

La (40)₂ vale anche con indici a, b scambiati. Similmente dalla (22)₃ si ha

$$(41) \quad p_a^* \cdot \dot{U} \cdot p_a^* = \dot{\lambda}_a, \quad p_a^* \cdot \dot{U} \cdot p_b^* = (\lambda_b - \lambda_a) \omega_{ab}^* .$$

Dalla (28), con le (19), (20), (22) si ottengono le relazioni

$$(42) \quad 2\dot{E}_o = g^\alpha g_\alpha^* \cdot \dot{C} \cdot g_\alpha^* g^\alpha, \quad \dot{C} = U \cdot \dot{U} + \dot{U} \cdot U, \\ g^\alpha g_\alpha^* = p_\alpha p_\alpha^* \cdot U^{-1} .$$

Quindi, con l'aiuto delle (22)₃ e (41) si hanno le componenti di $\dot{E}_o$

$$(43) \quad p_a^* \cdot \dot{E}_o \cdot p_a^* = \dot{\lambda}_a / \lambda_a, \quad p_a^* \cdot \dot{E}_o \cdot p_b^* = (\lambda_b^2 - \lambda_a^2) \omega_{ab}^* / 2 \lambda_a \lambda_b .$$

In base alle espressioni (40) e (43) la (37) si traduce nelle relazioni seguenti fra i tensori S_e e S :

$$(44) \quad (p_a^* \cdot S_e \cdot p_a^*) e_a' = (p_a \cdot S \cdot p_a) / \lambda_a \\ (p_a^* \cdot S_e \cdot p_b^*) (e_b - e_a) = (p_a \cdot S \cdot p_b) (\lambda_b^2 - \lambda_a^2) / 2 \lambda_a \lambda_b .$$

Si nota che in queste relazioni non interviene né λ_a né il tensore $\mathbf{R}^*$. Si controlla facilmente che, qualora si ponga $\mathbf{E}_e = \mathbf{E}$ e quindi $e_a = (\lambda_a^2 - 1)/2$ risulta $\mathbf{S}_e = \mathbf{S}''$: dunque questo tensore rientra fra quelli costruibili secondo Hill. Fra questi è particolarmente significativo quello che deriva dalla posizione $e_a = \ln \lambda_a$: in questo caso risultano esattamente eguali le componenti normali di $\mathbf{S}_e$ e $\mathbf{S}$ che entrano nella (44)₁, mentre le tangenziali risultano circa eguali per deformazioni non molto grandi.

Altro tensore allotropico, ma con diverso carattere dai precedenti è il «primo tensore di Piola-Kirchhoff»

$$(45) \quad \mathbf{T}' = s^\alpha \mathbf{g}_\alpha^* / j^*$$

nel quale le forze risultanti sulle facce del parallelepipedo attuale sono trasportate sulle corrispondenti facce iniziali. Un'importante proprietà di questo tensore si desume dalle (6), (7) che si possono trascrivere per la situazione sul parallelepipedo di lati $\mathbf{g}_\alpha^* d\xi^a$ considerando che si mantengono invariati gli incrementi d_a e le risultanti delle tensioni: quindi per la (7) si ha l'eguaglianza $\sigma A = \sigma^* A^*$. Dunque con questo tensore la (1)₁ è sostituita dalla relazione

$$(46) \quad \mathbf{T}' \cdot \mathbf{n}^* = \sigma A / A^*$$

che dà la tensione «nominale» (ossia riferita all'area iniziale). Le stesse considerazioni che portano a scrivere le (5)_{2,3} nelle coordinate ξ^α valgono per le condizioni di equilibrio dell'elemento nella configurazione iniziale se esso è soggetto ad una forza $j^* \mathbf{f}_v^* = j \mathbf{f}_v$ e ad una coppia $s^\alpha \times (\mathbf{g}_\alpha - \mathbf{g}_\alpha^*)$, ambedue riferite agli incrementi di coordinate ξ^α .

8. Formulazione differenziale

Alla configurazione attuale si imprime una variazione definita da spostamenti $\dot{\mathbf{x}} \delta t$ e si individuano gli effetti di un campo assegnato di «velocità» $\mathbf{v} = \dot{\mathbf{x}}$ mantenendo invariate le coordinate materiali sicché risulta $\dot{\mathbf{g}}_\alpha^* = 0$, $\dot{\mathbf{g}}_\alpha = \partial \mathbf{v} / \partial \xi^\alpha$. Nell'esame della configurazione spostata si assume per riferimento la attuale: ossia nelle relazioni in gioco, operata la derivazione rispetto al tempo, si sostituiscono ai dati di riferimento quelli attuali. Così dal tensore (19) si deduce il gradiente di velocità

$$(47) \quad \dot{\mathbf{F}}_o = \dot{\mathbf{g}}_\alpha \mathbf{g}^\alpha.$$

Da questo si separa la parte simmetrica

$$\dot{\mathbf{E}}_o = (\dot{\mathbf{F}}_o + \dot{\mathbf{F}}_{or})/2$$

che misura la velocità di deformazione in accordo con la (28), e la anti-simmetrica, velocità di rotazione

$$(48) \quad \mathbf{W} = (\dot{\mathbf{F}}_o - \dot{\mathbf{F}}_{or})/2 .$$

Le due velocità sono misurate rispetto alla configurazione attuale. La componente cartesiana del tensore (47), essendo $\mathbf{g}^\alpha \cdot \mathbf{i}_\mu = \partial \xi^\alpha / \partial y^\mu$, è

$$(49) \quad \mathbf{i}_\lambda \cdot \dot{\mathbf{F}}_o \cdot \mathbf{i}_\mu = \partial v_\lambda / \partial y^\mu .$$

Dunque $\dot{\mathbf{E}}_o$ si allaccia all'usuale forma lineare della deformazione, mentre $\mathbf{W}$ si collega alla rotazione. Sostituendo nella (48) l'espressione (22)₁ e ponendo $\mathbf{R} = \mathbf{U} = \mathbf{I}$, tensore unitario, si ottiene $\mathbf{W} = \mathbf{R}_o$ e perciò

$$(50) \quad \dot{\mathbf{F}}_o = \dot{\mathbf{E}}_o + \dot{\mathbf{R}}_o .$$

Si osserva che la componente $\mathbf{i}_c \cdot \mathbf{W} \cdot \mathbf{i}_b$ rappresenta la proiezione su $\mathbf{i}_a$ della velocità angolare.

Per il tensore $\mathbf{T}$ nella forma (15)₁, operando come per la (39), si ha

$$(51) \quad \dot{\mathbf{T}} = \dot{\mathbf{F}}_o \cdot \mathbf{T} + \mathbf{g}_\alpha \dot{t}^{\alpha\beta} \mathbf{g}_\beta + \mathbf{T} \cdot \dot{\mathbf{F}}_{or}$$

Si definisce la «derivata di Jaumann $\mathbf{T}^o$ » di $\mathbf{T}$ con la proprietà:

$$(52) \quad \mathbf{T}^o \cdot \mathbf{n} \equiv \dot{\boldsymbol{\sigma}} - \dot{\mathbf{R}}_o \cdot \boldsymbol{\sigma} .$$

Operando sul tensore $\mathbf{n}$, essa dà la variazione di tensione misurata rispetto alla particella, come differenza fra la variazione assoluta e quella di trascinamento $\dot{\mathbf{R}}_o \cdot \boldsymbol{\sigma}$ dovuta alla velocità rotatoria della particella. Sostituendo le espressioni $\dot{\boldsymbol{\sigma}} = \dot{\mathbf{T}} \cdot \mathbf{n} + \mathbf{T} \cdot \dot{\mathbf{n}}$, $\dot{\mathbf{n}} = \dot{\mathbf{R}}_o \cdot \mathbf{n}$, per l'arbitrarietà della scelta di $\mathbf{n}$, risulta mediante le (52), (51)

$$(53) \quad \mathbf{T}^o = \mathbf{g}_\alpha \dot{t}^{\alpha\beta} \mathbf{g}_\beta + \dot{\mathbf{E}}_o \cdot \mathbf{T} + \mathbf{T} \cdot \dot{\mathbf{E}}_o = \dot{\mathbf{T}} + \mathbf{T} \cdot \mathbf{W} - \mathbf{W} \cdot \mathbf{T} .$$

Moltiplicando per $\mathbf{g}^\alpha$, $\mathbf{g}^\beta$ primo e secondo membro si ha

$$(54) \quad \mathbf{g}^\alpha \cdot \mathbf{T}^o \cdot \mathbf{g}^\beta = \dot{t}^{\alpha\beta} + \mathbf{g}^{\alpha\lambda} \dot{\gamma}_{\lambda\mu} t^{\mu\beta} + t^{\alpha\lambda} \dot{\gamma}_{\lambda\mu} \mathbf{g}^{\mu\beta} .$$

Qualora si adoperino le componenti definite dalla (17), derivando l'eguaglianza $s_n^\alpha{}_\beta = s^{\alpha\lambda} g_\lambda \cdot i_\beta$, posto che le coordinate ξ^α coincidano nel riferimento con le y^α , si ottiene

$$(55) \quad \dot{s}_n^\alpha{}_\beta = \dot{s}^\alpha{}_\beta + s^{\alpha\lambda} \partial v_\beta / \partial y^\lambda .$$

La relazione fra le quantità $\dot{t}^{\alpha\beta}$ e $\dot{s}^{\alpha\beta}$ si deduce considerando che

$$(56) \quad \dot{j} = \partial v_\alpha / \partial y^\alpha .$$

La relazione fra incrementi di tensione e deformazione viene scritta

$$(57) \quad g^\alpha \cdot T^\sigma \cdot g^\beta = C^{\alpha\beta\gamma\delta} \dot{\gamma}_{\gamma\delta} .$$

I coefficienti $C^{\alpha\beta\gamma\delta}$ dipendono in generale non solo dalla storia del processo di deformazione ma anche dai rapporti attuali fra le velocità di deformazione: questa seconda dipendenza introduce una nonlinearietà nelle velocità. Essa manca per il corpo elastico e anche per l'elasto-plastico in un definito campo di variazione dei detti rapporti per il quale non si presentano deformazioni anelastiche. In questo, se può ammettersi la reversibilità del ciclo elastico, la matrice 6×6 , nella quale possono disporsi i coefficienti in virtù della simmetria delle componenti di tensione e deformazione, risulta simmetrica. Nella (57) sono introdotte le derivate di Jaumann per dare carattere «oggettivo» alla relazione, indipendente dal moto rigido della terna p_α . Si nota, secondo Hill, che è parimenti legittimo porre a 1° membro le derivate $\dot{t}^{\alpha\beta}$ o $\dot{s}^{\alpha\beta}$ poichè quelle componenti sono misurate su una base che, rispetto alla p_α , subisce modifiche dell'ordine delle deformazioni. Altrimenti può vedersi ciò dalla (54): la differenza fra la derivata di Jaumann e la sostanziale delle tensioni è costituita da termini che possono portarsi a 2° membro della (57), conglobandone gli effetti nei coefficienti C .

Il sistema nelle velocità v va completato con le condizioni di equilibrio che possono scriversi adoperando l'equazione dei lavori

$$(58) \quad \dot{L}_e = \int \dot{s}^\alpha \cdot (\partial \delta v / \partial \xi^\alpha) d\xi^1 d\xi^2 d\xi^3 .$$

Si noti che il campo di integrazione non cambia nel tempo: quindi con le stesse variabili di integrazione, la velocità di variazione del lavoro esterno per la parte che concerne i carichi agenti sul volume del corpo si ottiene con l'integrando $\delta v \cdot (j f_v)^*$.

9. Considerazioni comparative

Prodotti scalari. Poiché nella scrittura esplicita le operazioni algebriche sono ricondotte nell'ambito vettoriale, coerentemente i prodotti (scalari) sono sempre segnati con punto, anche nella scrittura implicita di tensori e vettori. Così si fa, ad esempio nella (20): nel prodotto delle diadiche le basi contigue danno una matrice 3×3 . Matrici dello stesso tipo si riscontrano nei cambiamenti di base (§ 3). Peraltro nella letteratura, più frequentemente, per il prodotto fra tensori o fra questi e vettori non si pone alcun segno fra i due simboli. Allora, per il prodotto tensoriale (in particolare fra le due basi di un tensore) deve disporsi il segno di diade $\otimes$.

Il doppio prodotto fra due tensori $\mathbf{A}$, $\mathbf{B}$ di componenti $a^{\alpha\beta}$, $b_{\alpha\beta}$ su basi reciproche, ossia l'espressione $a^{\alpha\beta} b_{\alpha\beta}$ (oppure $a^{\alpha\beta} b_{\beta\alpha}$), in linguaggio tensoriale si designa come traccia del prodotto di $\mathbf{A}$ per $\mathbf{B}_\tau$ (oppure per $\mathbf{B}$). La notazione in uso $\mathbf{A}:\mathbf{B}$ (oppure $\mathbf{A}.. \mathbf{B}$) è stata qui adottata.

Derivate covarianti. L'equazione (1)₂ in termini di componenti è scritta

$$(59) \quad t_h^{\alpha\beta}{}_{,\alpha} + \mathbf{f}_v \cdot \mathbf{h}^\beta = 0.$$

Il primo termine è detto «derivata covariante di tensore controvariante»: sviluppato secondo l'apposito formulario dà la stessa espressione che si ricava alla (5)₂ moltiplicata per $\mathbf{h}^\beta/j_h$ se nella (9) si pone $\mathbf{h}' = \mathbf{h}$. Il calcolo diretto dà maggiori possibilità di scelta e per la terna $\mathbf{h}'$ non esige la determinazione della rete coordinata dalla quale quella terna è derivata. Si nota che la giustificazione delle (5)_{2,3} come quella della (25), si allaccia a concetti familiari nei corsi di Ingegneria.

Prova su modello. Il modello, poniamo un parallelepipedo di spigoli $\mathbf{g}_{s\alpha}$, sia deformato nel tempo in modo che i prodotti $\mathbf{g}_{s\alpha} \cdot \mathbf{g}_{s\beta}$ assumano gli stessi valori $g_{\alpha\beta}$ della terna naturale in P ; sulle sue facce si susseguiranno le stesse tensioni $t^{\alpha\beta}$ che si hanno in P purché la coincidenza sia esistita in un tempo abbastanza remoto (o il corpo sia stato scarico nella configurazione iniziale): ciò si verifica indipendentemente dai moti rigidi che il modello possa avere effettuato. Con riferimento alle direzioni principali della deformazione che porta dallo stato iniziale all'attuale la deformazione si riduce a una pura similitudine. Pertanto in [1] p. 66, è enunciato il «teorema fondamentale» per il quale la risposta a qualsiasi storia di deformazioni è nota se si conosce la risposta

«to pure stretch histories». Questo non deve indurre all'errore di ritenere che il comportamento sia definito dalla storia di 3 parametri, le dilatazioni principali: la formulazione indicata mostra che i parametri influenti sono 6; come tali si possono assumere le dette dilatazioni e i 3 parametri necessari per definire l'orientamento della terna principale nel corpo.

Leggi costitutive olonome. Per il corpo ideale elastico, nota la configurazione iniziale, basta la attuale per conoscere lo stato di tensione. Per la formulazione più conveniente di una legge costitutiva olonoma, nella quale cioè entrino i valori attuali della deformazione Hill [6, 7] suggerì la costruzione di tensori «associati» S_e , E_e posti sulle stesse basi (iniziali) del tensore delle deformazioni finite E per formulare la dipendenza funzionale

$$(60) \quad S_e = f(E_e) .$$

Al tensore E_e si assegnano le dilatazioni principali come funzioni di quelle di E . Ad una variazione generica della configurazione attuale corrisponde una variazione nota δE_e : eguagliando il lavoro delle tensioni S_e per questa variazione a quello delle tensioni S per la variazione di configurazione attuale si stabiliscono 6 relazioni fra le componenti dei due tensori di tensione; sicché se S_e è noto dalla (60) si ottengono le componenti di S .

Relazioni dalla forma (60) sono adoperate per la risoluzione di problemi di grandi deformazioni secondo la formulazione detta «lagrangiana aggiornata (updated)»: si procede per intervalli di tempo relativamente ampi, assumendo in ciascun passo la configurazione iniziale di esso per riferimento. Usualmente la relazione (60) viene posta fra i tensori S'' e E : così, ad esempio in [8]. È stata talora usata [9] la misura di deformazione $e_\alpha = \ln \lambda_\alpha$ per la quale il tensore S viene ad avere sulla terna p_α le stesse componenti di S_e su p_α^* , per deformazioni non troppo grandi: questa concordanza dà alla legge costitutiva un carattere meno ipotetico. La misura logaritmica della deformazione, con i suoi valori crescenti da 0 a ∞ con gli allungamenti e decrescenti da 0 a $-\infty$, con gli accorciamenti è più naturale che quella data da $(\lambda_\alpha^2 - 1)/2$.

Leggi costitutive incrementalì. La forma incrementale (57) delle leggi costitutive può dare adito ad una rappresentazione più rispondente al reale comportamento del materiale con coefficienti C opportunamente individuati in ciascun caso. Per la risoluzione di problemi di grandi de-

formazioni in generale si deve seguire il processo di carico per passi più o meno ridotti secondo la rapidità di variazione dei coefficienti. Lungo il percorso, le situazioni di biforcazione e di instabilità sono individuabili secondo Hill [7]. Questo autore e altri sullo stesso indirizzo [5] fanno ricorso alle tensioni nominali, i cui incrementi possono determinarsi mediante il tensore

$$(61) \quad \mathbf{T}' = \dot{s}_n^{\alpha\beta} \mathbf{i}^\beta \mathbf{i}_\alpha^*$$

in riferimento cartesiano, con la seconda terna, di sezionamento, nella configurazione iniziale e quindi invariata come la prima nell'intervallo dt . L'equazione dei lavori virtuali è scritta nella forma derivata (58) dove le componenti $\dot{s}_n^{\alpha\beta}$ vengono espresse mediante le (54), (55) in funzione delle velocità di deformazione e delle derivate di Jaumann delle tensioni, che le (57) mettono in connessione con le dette velocità. Si osserva ancora che nel calcolo di $\dot{L}_e$, la semplice espressione che per la variazione delle forze di superficie e di volume si ha in termini di tensioni nominali non esime dal calcolo delle variazioni delle forze sugli elementi in deformazione.

Si osserva infine che, mentre le difficoltà di impostazione possono essere ridotte da una chiara formulazione e poi vinte dalla traduzione opportuna del problema analitico in una indagine numerica, le incertezze sulle basi fisiche, in particolare, sulla determinazione dei coefficienti C nelle (57) non svaniscono sinché non si disponga di adeguate informazioni sul comportamento del materiale.

RIFERIMENTI

- [1] TRUESDELL C., NOLL W., *The nonlinear field theories of Mechanics*, Encyclopedia of physics, Vol. III/3, Springer, Berlin, 1965.
- [2] MALVERN L.E., *Introduction to the Mechanics of a continuum medium*, Prentice-Hall, Englewood Cliffs, N.J., 1969.
- [3] ALGOSTINO F., CICALA P., *Sull'analisi delle grandi deformazioni*, Atti Acc. Sc. Torino, Vol. 122, 1988, pp. 126-138.
- [4] ALGOSTINO F., *Approccio vettoriale all'analisi delle grandi deformazioni e relative tensioni*, Mem. Acc. Sc. Torino, Serie V, Vol. 12, 1988, pp. 5-45.
- [5] MCMEEKING R.M., RICE J.R., *Finite-element formulations for problems of large elastic-plastic deformation*. Int. J. Solids Structures Vol. 11, 1975, pp. 601-616.
- [6] HILL R., *On constitutive inequalities for simple materials*, J. Mech. Phys. Solids, Vol. 15, 1967, pp. 79, Vol. 16, 1968, pp. 315-322.
- [7] HILL R., *Aspects of invariance in solid mechanics*, Advances in Appl. Mech., Vol. 18, 1978, pp. 1-75.
- [8] NAGTEGAAL J.C., DE JONG J.E., *Some computational aspects of elastic-plastic large strain analysis*. Int. J. Num. Methods Engin., 1981, pp. 15-41.
- [9] BATHE K.J., KOJC M., WALCZAK J., *Some developments in methods for large strain elasto-plastic analysis*. Proc. Conference Computational Plasticity, Barcelona, April 1987.

Ottimizzazione della configurazione di un moderno velivolo a breve raggio

Piero PELAGALLI*

Introduzione

Questa nota si propone di presentare una sintesi di alcuni risultati degli studi di progettazione preliminare eseguiti, nell'ambito di un accordo internazionale, dal Gruppo Velivoli da Trasporto della Società Alenia su una nuova famiglia di velivoli a breve raggio con capacità compresa fra 90 e 130 posti circa. La Fig. 1 mostra una vista d'insieme del velivolo maggiore in una delle configurazioni studiate.

Il processo che abbiamo chiamato di «ottimizzazione», inteso a definire la configurazione generale dei velivoli citati, comprende numerosi cicli di progettazione, analisi economica e valutazione operativa da parte delle compagnie aeree interessate. Si comprende perciò che si tratta di un processo molto complesso, di cui in questa breve esposizione non è possibile dare se non un'idea schematica e approssimativa. Gli obiettivi, che in esso si perseguono, sono sia di qualità del prodotto sia di tipo economico. Essenzialmente si tratta di definire un prodotto «di qualità», nel senso che soddisfi le esigenze dei Clienti (aerolinee e passeggeri) sostanzialmente meglio della competizione, e che al tempo stesso possa essere utilizzato con risultati economici positivi dalle aerolinee ed essere prodotto dai costruttori realizzando un utile.

Raggiungere questi obiettivi non è affatto facile nella attuale situazione di mercato, estremamente competitiva, sia per le aerolinee che per i costruttori. Di qui la necessità di una lunga e paziente attività di ottimizzazione, cioè di una approfondita analisi del ritorno economico di tutte le risorse tecnologiche disponibili e di una valutazione comparativa di tutte le opzioni progettuali escogitabili per scegliere quelle che consentano di soddisfare i requisiti operativi al minore costo possibile. Questa attività deve proseguire oltre la fase di progettazione preliminare, oggetto di questa nota, andando ad interessare via via aspetti più dettagliati di definizione del prodotto.

* ALENIA Aeronautica, Pomigliano d'Arco, Napoli.



FIG. 1

1. Obiettivi della progettazione preliminare

Scopo della fase di progettazione preliminare è di definire:

- i requisiti generali di progetto (carico pagante, autonomia, prestazioni, etc.);
- la configurazione generale del velivolo (superficie e geometria dell'ala e degli impennaggi, sezione trasversale e forma della fusoliera, tipo e posizione dei motori, architettura generale della cabina di pilotaggio, dell'interno cabina passeggeri, sistemazione degli equipaggiamenti principali, etc.);
- il progetto aerodinamico preliminare (con particolare riguardo all'ala);
- il sistema propulsivo, in collaborazione con i motoristi, in modo da sposare in modo ottimale i requisiti di progetto (si intende che si parla di un motore nuovo da svilupparsi insieme al velivolo);
- i principali sottosistemi (avionica, comandi di volo, condizionamento, elettrico, idraulico, etc.);
- la tecnologia dei materiali strutturali (identificando i benefici potenziali, i costi e il rischio tecnico);
- le masse ad un livello sufficiente di dettaglio e precisione;
- i costi (costo del velivolo, costi di sviluppo, costi operativi).

Si tratta, come sopra accennato di un processo iterativo (in Fig. 2 un tentativo di rappresentazione schematica), che deve coinvolgere anche le aerolinee interessate. Si sottolinea l'importanza di ciò: al fine di arrivare ad un prodotto che risponda alle esigenze del mercato, occorre che gli attori primari del mercato, cioè le aerolinee, entrino fin dall'inizio nella definizione del prodotto. In particolare, è indispensabile che esse intervengano negli inevitabili compromessi in termini di prestazioni e nelle scelte progettuali, che i vincoli di costo impongono.

Fra i più importanti di tali vincoli si deve citare la necessità di utilizzare il concetto di «FAMIGLIA» di velivoli, fin dall'inizio della progettazione preliminare. Infatti a causa degli ingenti investimenti richiesti per lo sviluppo di un nuovo velivolo, è oggi impensabile di sviluppare ed immettere sul mercato un prodotto singolo. La via obbligata è di concepire dall'inizio un gruppo («famiglia») di velivoli con la massima comunanza fra di loro, in grado di coprire un'area di mercato quanto più vasta possibile.

In questo modo si possono ripartire su più prodotti i costi di sviluppo, gli investimenti industriali e commerciali, le spese per l'assistenza etc.

Anche dopo lo sviluppo dei membri iniziali della famiglia, i costruttori debbono sforzarsi di utilizzare al massimo gli investimenti già fatti adottando i prodotti alle richieste del mercato o derivando da essi nuovi prodotti atti a soddisfare nuove esigenze.

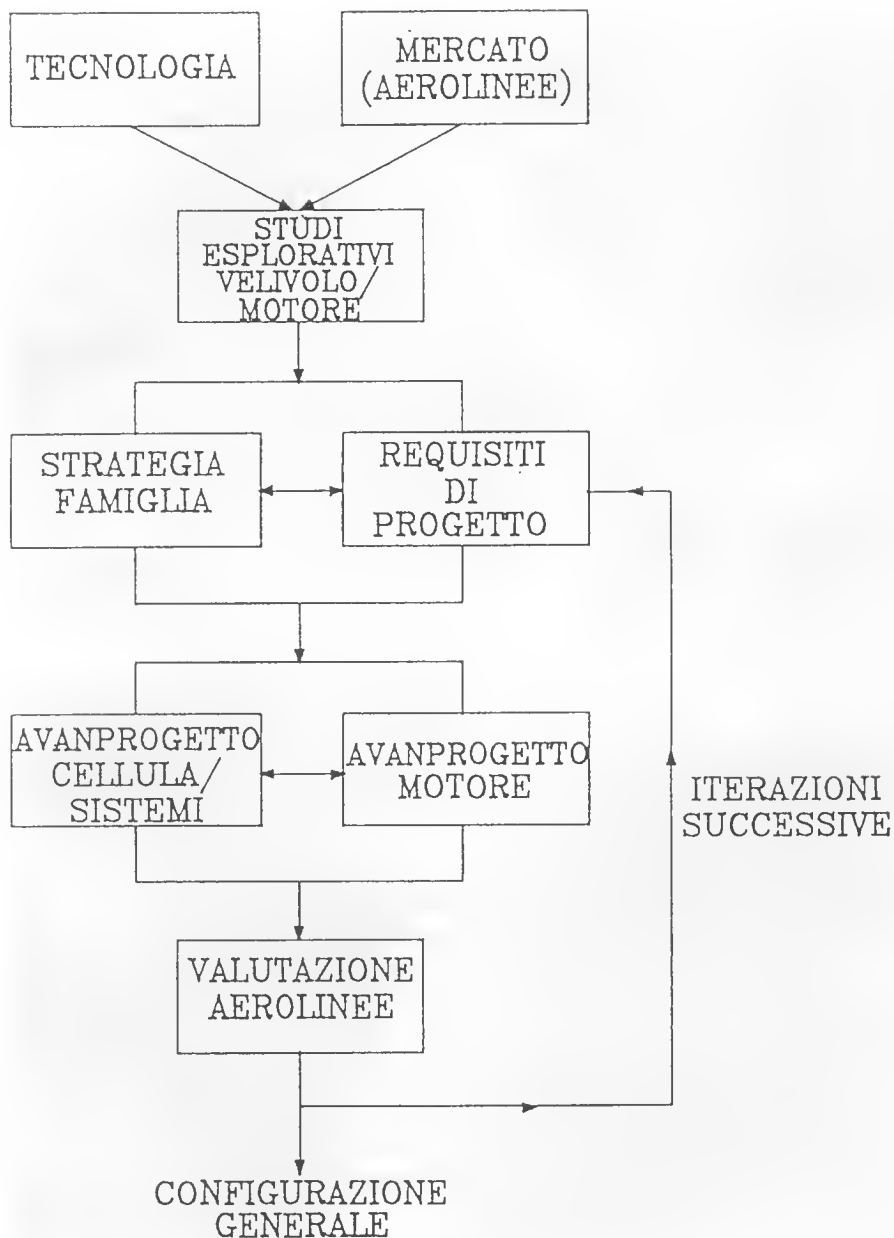


Figura 2 - Schema del processo di progettazione preliminare.

2. Criteri di valutazione

In fase di progettazione preliminare occorrono strumenti di valutazione semplici, che consentano di prendere rapidamente una decisione negli studi, tipici di questa fase, sulle scelte tecniche fondamentali. Il metodo più utilizzato si basa sulla valutazione dell'effetto di ogni scelta tecnica sul «costo operativo diretto» (si userà nel seguito la sigla internazionale D.O.C., Direct Operating Cost), il cui significato è ben noto.

È ovvio infatti che la validità economica e la competitività sul mercato di un velivolo da trasporto commerciale sono tanto migliori quanto più basso è il costo globale di esercizio (total operating cost) sostenuto dall'aerolinea per l'impiego operativo della macchina. Di tale costo, come noto, il D.O.C. rappresenta la parte comprendente le voci di costo, che derivano direttamente dalla utilizzazione del velivolo e dipendono dalle sue caratteristiche tecniche ed economiche. Tali voci sono essenzialmente ed in ordine di importanza:

- oneri finanziari (ammortamento del capitale investito nel velivolo, assicurazione);
- spese per l'equipaggio;
- costi di manutenzione;
- diritti aeroportuali e di navigazione;
- costo del combustibile e dei lubrificanti consumati.

In Fig. 3 viene riportato un esempio tipico di ripartizione di tali costi, nonché una formulazione semplificata del D.O.C. Non è questo il luogo per discutere i metodi per il calcolo del D.O.C. Importa solo ricordare che vengono usati molti metodi diversi, anche se non mancano gli sforzi di unificazione (ad es., il metodo ATA). In pratica, ogni grande costruttore, ogni grande aerolinea o associazione di aerolinee utilizza un proprio metodo. In particolare, nel caso delle aerolinee, ognuna utilizza formule «personalizzate», che tengono conto della politica degli ammortamenti, della composizione della flotta, delle pratiche di manutenzione, della struttura della rete, dei contratti col personale etc., particolari di quella aerolinea. Non si possono quindi in generale confrontare valori di D.O.C. di origine diversa e, comunque, ciò che è significativo non è tanto il valore assoluto del D.O.C., ma le variazioni di esso in funzione delle caratteristiche del velivolo.

Importa anche osservare che il criterio del D.O.C. è valido in sede di progettazione preliminare, ma è troppo grossolano per poter essere utilizzato nelle fasi di definizione e di progettazione esecutiva. In queste sono più appropriati criteri più precisi e capaci di utilizzare dati di maggior dettaglio, come quello del «life cycle cost».

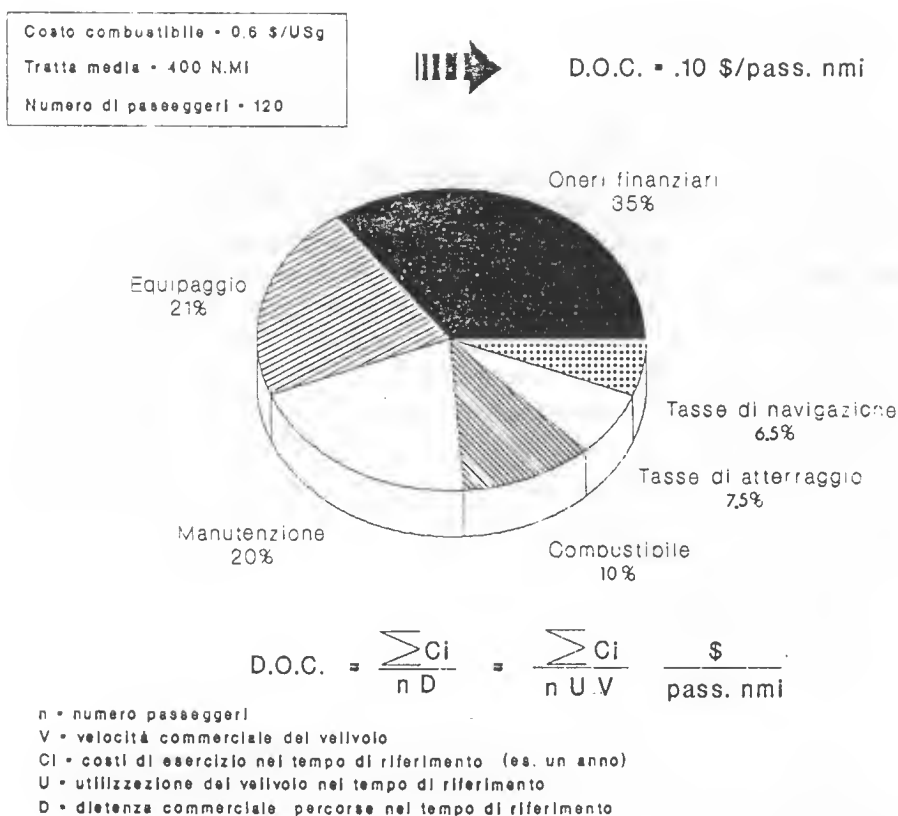


Figura 3 - Esempio di ripartizione del D.O.C. (Costo Operativo Diretto).

3. Strumenti di progettazione preliminare

In fase di progettazione preliminare è essenziale disporre di strumenti di calcolo che siano in grado di fornire risultati in tempi brevi e con buona approssimazione.

Nello studio in discussione è stato utilizzato un complesso di codici di calcolo, denominato SCAN, sviluppato dalla Direzione Tecnica del Gruppo Aerei da Trasporto dell'Aeritalia, per effettuare il dimensionamento rapido di un velivolo da trasporto. Tali codici permettono anche di valutare in modo sintetico ed efficiente l'effetto che i requisiti hanno sulla configurazione generale del velivolo, nonché di valutare le prestazioni del velivolo e le sue caratteristiche economiche (costi ricorrenti

e non ricorrenti del programma, costo del velivolo, costi operativi). Lo SCAN viene inoltre utilizzato per studi di sensibilità, per confrontare configurazioni diverse, per valutare l'effetto di tecnologie alternative (ad es. materiali compositi) sulle prestazioni e sulle caratteristiche economiche del velivolo.

La struttura dello SCAN (Fig. 4) è costituita da alcuni «macromoduli» (GEOMETRIA, AERODINAMICA, PROPULSIONE, ANALISI DI MISSIONE, PESI E CENTRAMENTO, ECONOMIA), integrati da un programma principale (MAIN). Questo fa sì che, durante la procedura di dimensionamento del velivolo, vengano tenuti in conto tutti gli effetti derivanti dall'interazione fra le differenti variabili di progetto. In Fig. 5 è riportato un esempio di applicazione dello SCAN, mentre di seguito viene riportata una breve descrizione dei moduli.

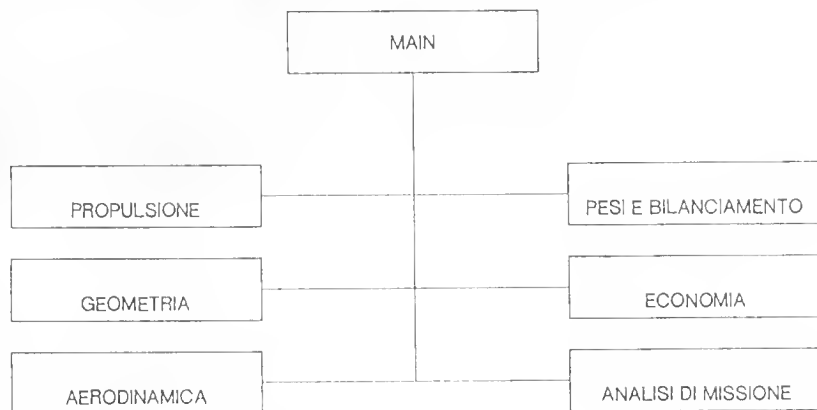


Figura 4 - Struttura programma «S.C.A.N».

Modulo geometria

Dati di input: numero passeggeri, caratteristiche geometriche o dimensionali di ala e impennaggi.

Dati output: lunghezza e diametro fusoliera, aree di ala e impennaggi, aree bagnate etc.

Modulo aerodinamica

Dati di input: caratteristiche profili, geometria, condizioni di volo, tipo di ipersostentatori.

Dati output: polari di alta velocità, polari di bassa velocità nelle configurazioni «clean», di decollo e di atterraggio.

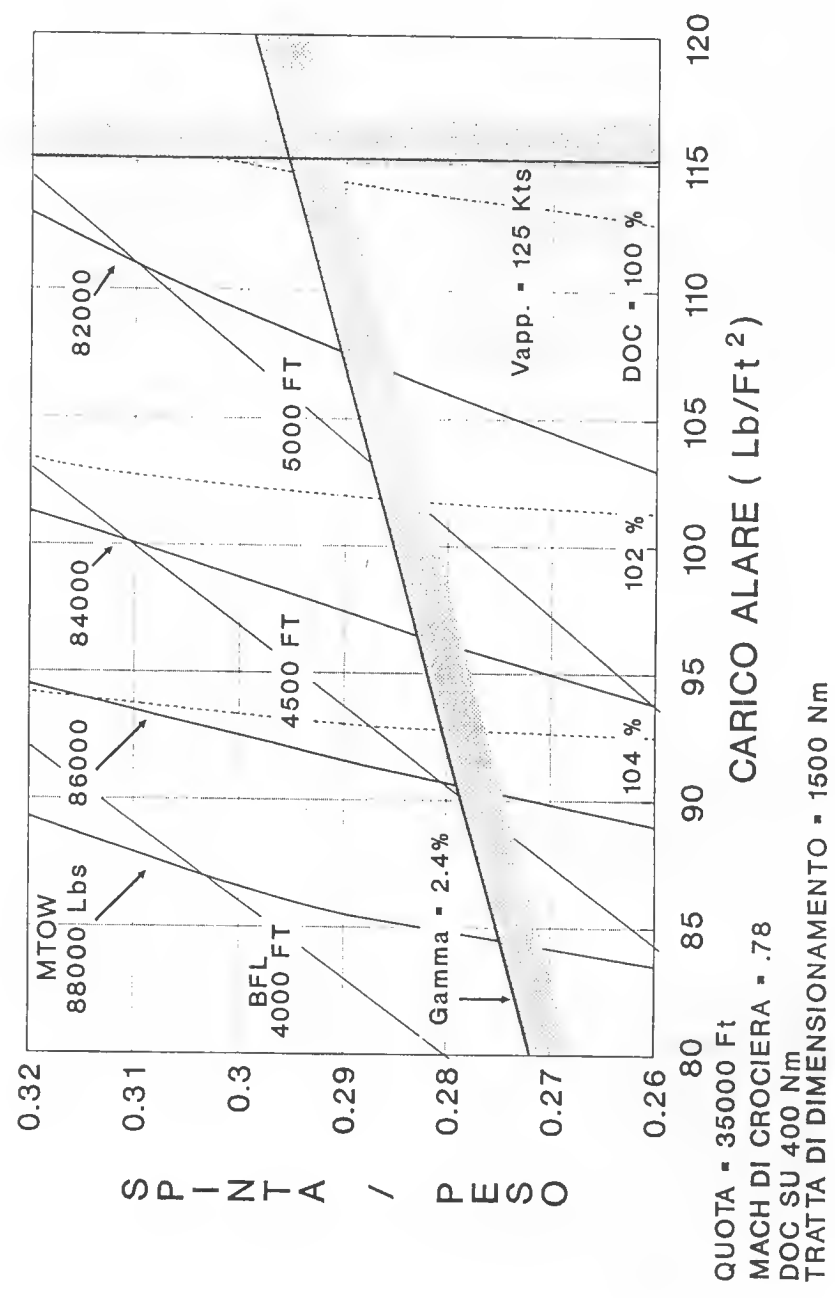


Figura 5 - Esempio di applicazione del programma SCAN velivolo da 90 posti.

Modulo propulsione

Dati di input: ciclo motore, caratteristiche dimensionali e inerziali del motore di riferimento in funzione della spinta.

Dati output: dimensionamento del motore e sue caratteristiche.

Modulo pesi e centramento

Dati di input: geometria, pesi di missione.

Dati output: pesi di progetto, peso di struttura ed impianti secondo ATA, posizionamento ala per ottenere il centramento voluto.

Modulo analisi di missione

Questo modulo analizza le varie fasi della missione del velivolo: taxi, decollo salita, crociera, discesa, atterraggio. Calcola anche la lunghezza di pista bilanciata, la velocità massima, etc. Quando è specificato il requisito di autonomia, il programma dimensiona il velivolo, che soddisfa tale requisito.

Modulo economia

Dati di input: pesi, motorizzazione, consumi combustibili etc. (dagli altri moduli).

Dati output: costi di sviluppo, costi ricorrenti, prezzo del velivolo, costi operativi.

4. Requisiti di missione

I requisiti di missione hanno una influenza molto importante sulle caratteristiche economiche del velivolo. Le scelte al riguardo debbono essere fatte dalle aerolinee. Il compito dei costruttori è soprattutto quello di dare ad esse tutte le informazioni necessarie a comprendere le conseguenze delle loro scelte. Quando non si segue questa linea, si rischia un probabile fallimento (si ricorda il caso Mercure). Non è ovviamente possibile in questa sede una discussione esauriente sull'argomento, per cui ci si limiterà ad un breve esame dell'influenza dei requisiti più significativi.

Carico pagante. Diverse indagini con domande poste direttamente ad un campione rappresentativo (circa il 70% del totale) delle aerolinee

interessate a questo tipo di velivolo, ha portato alla conclusione che la richiesta si concentra su due distinti valori di capacità di carico:

- 85-90 posti
- 120-130 posti

Si è potuto quindi concludere che occorre sviluppare due tipi di velivoli, uno più piccolo ed uno più grande. Studi di dettaglio hanno mostrato che è possibile realizzare due velivoli da 90 e 125 posti circa con la stessa superficie alare, stessi impennaggi, stessa gondola motore, motore con spinta aumentata variando il ciclo termodinamico e minime modifiche fisiche. Le altre varianti del velivolo maggiore sono un sistema di ipersostentatori più sofisticato, rinforzi strutturali, carrello e ruote rinforzati e freni potenziati, come richiesto dai pesi di progetto più elevati. A parte queste modifiche, il velivolo da 125 posti viene derivato da quello da 90 posti semplicemente aumentando la lunghezza della fusoliera. In questo modo si possono utilizzare per i due velivoli le stesse attrezzature di costruzione ed una elevata percentuale delle parti, realizzando così una «famiglia» iniziale di prodotti, fatto di grande importanza economica, come sopra discusso. Inoltre il progetto iniziale deve permettere la possibilità di sviluppare altri velivoli di maggiore capacità e autonomia utilizzando una nuova ala.

Autonomia. Le indagini presso le aerolinee hanno evidenziato che un'autonomia di progetto di 1500 n. mi (miglia nautiche) è sufficiente per la maggior parte delle tratte previste per questo tipo di velivolo. Volendo coprire il 100% di tali tratte occorrerebbe una autonomia di oltre 2000 n. mi. In questo modo, tuttavia, il velivolo risulterebbe più pesante e costoso. Il DOC nella tipica tratta media (circa 400 n. mi) ne risulterebbe pertanto penalizzato, come evidenziato in Fig. 6 ricavata col programma SCAN. È stato deciso di mantenere un'autonomia di 1500 n. mi, rinviandone l'aumento dopo la prima certificazione, allorché sarà quasi certamente possibile aumentare le masse di progetto sulla base dei margini, che le prove strutturali normalmente evidenziano sulla cellula e sul carrello.

Velocità massima. Influenza direttamente la geometria dell'ala (angolo freccia, spessore, allungamento), la spinta motore (a meno che il motore non venga dimensionato dal decollo dalla salita) e il dimensionamento strutturale. Si può dire, quindi, che all'aumentare della velocità il velivolo diventa più pesante e costoso. Tuttavia, la velocità compare a denominatore nella formula del DOC (aumenta la produttività del velivolo). L'effetto risultante è un andamento abbastanza piatto del DOC

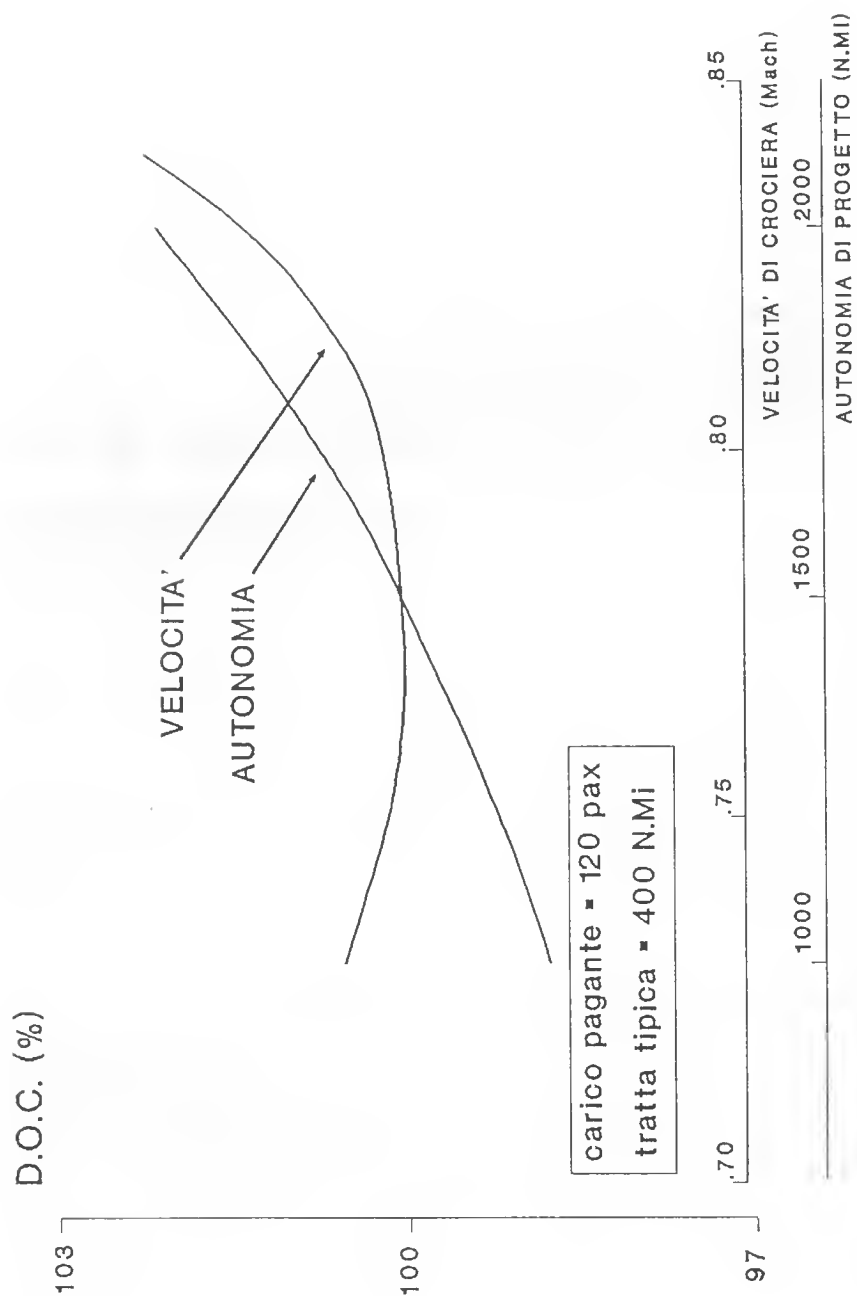


Figura 6 - Effetto sul DOC dei requisiti di autonomia e di velocità.

nel campo di velocità numero di Mach 0,75-0,80, come mostrato in Fig. 6. Si è pertanto scelta una velocità di progetto di 0,81 Mach essenzialmente per motivi di competitività.

Prestazioni di decollo e atterraggio. Per velivoli del tipo in esame la superficie alare è determinata dal requisito di distanza di atterraggio (equivalente a quello di velocità di avvicinamento fissata a 125 kts), ovviamente per il velivolo maggiore. Poiché anche la dimensione degli impennaggi è direttamente proporzionale a quella dell'ala, si può dire che questo requisito influenza direttamente il 20% del peso del velivolo.

Invece, come sopra accennato, le prestazioni di decollo (distanza di decollo bilanciata e gradiente di salita in monomotore) possono determinare il dimensionamento del motore. Nel caso in esame ciò si verifica per il 125 posti, ma non per il 90 posti.

L'indagine presso gli utilizzatori ed un esame dettagliato della lunghezza delle piste potenzialmente utilizzabili da questo tipo di velivolo, hanno evidenziato che sarebbe desiderabile potere operare da piste da 4500 piedi (circa 1370 m).

Tuttavia, vista la penalità che un requisito di questo genere imporrebbe sul DOC (Fig. 7) e tenuto conto dei risultati di un'analisi puntuale sulle tratte (e quindi sui pesi) di maggiore interesse in partenza dai vari aeroporti, è stato deciso di rilassare il requisito a 5000 piedi di lunghezza di pista al peso massimo al decollo. In questo modo si è ottenuto di limitare la penalità sul DOC, pur conservando la capacità di operare dalla maggior parte degli aeroporti di interesse e mantenendo un livello di prestazioni superiore alla concorrenza.

Da notare, infine, che la scelta di elevate prestazioni di decollo e atterraggio rende non critica (a differenza di molti altri bimotori) la tangenza in monomotore, che per entrambi i velivoli risulta superiore a 18000 piedi, e, quindi, largamente sufficiente.

5. Definizione di alcuni elementi della configurazione

La definizione della configurazione del velivolo comporta lo studio di un numero grandissimo di aspetti, che è impossibile riassumere in uno spazio così ristretto. Ci si limiterà quindi ad alcuni esempi fra i più significativi relativi alla configurazione generale, senza menzionare studi anche molto importanti relativi alle strutture e ai sistemi.

Sistema propulsivo. Data per scontata la formula bimotore, per un velivolo destinato a rimanere in servizio ben oltre l'inizio del prossimo

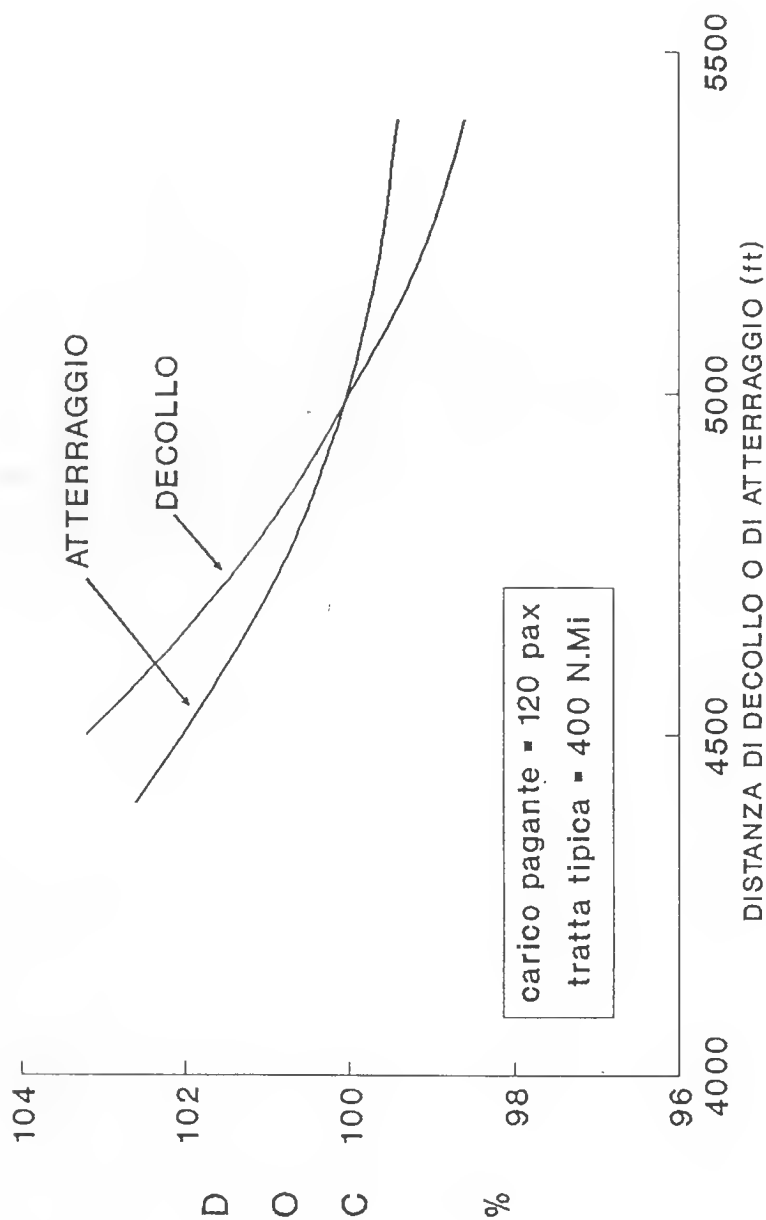


Figura 7 - Effetto sul DOC dei requisiti di decollo e atterraggio.

secolo e che utilizza un motore sviluppato appositamente, si impone uno studio sul tipo migliore di motorizzazione, considerando anche soluzioni non convenzionali.

L'unica grande innovazione tecnologica verificatasi nel campo della propulsione nell'ultimo decennio è stata lo sviluppo e la riuscita sperimentazione del motore con eliche ad alta velocità (Fig. 9), designato come «propfan» o «unducted fan» (General Electric). Il vantaggio potenziale di questo tipo di motore in termini di rendimento propulsivo è notevole (Fig. 8). Teoricamente, nel campo di velocità di interesse esso permette una riduzione del consumo di combustibile di oltre il 20% rispetto ai moderni motori turbofan ad alto rapporto di diluizione.

Allo scopo di determinare in modo realistico il vantaggio di questo motore una volta installato su un velivolo, sono state progettate due configurazioni secondo gli stessi requisiti, una con motori propfan e l'altra con motori turbofan (Fig. 10 e 11). Ne è risultato un vantaggio per il propfan inferiore al previsto (Fig. 12), malgrado si sia verificata una notevole riduzione di consumo (circa del 30%) e quindi anche del peso al decollo (6%) con questo propulsore. I motivi di ciò sono la limitata incidenza del consumo di combustibile sul D.O.C. (Fig. 3), ai costi attuali del combustibile, il più elevato costo del motore propfan e di alcuni sistemi velivolo con questo motore e più alti costi di manutenzione dovuti alla maggiore complessità di questo propulsore. È solo con costi molto elevati del combustibile, al momento non prevedibili, che il propfan presenta un significativo vantaggio economico (Fig. 13).

Ovviamente questi risultati hanno fatto cadere qualsiasi interesse al propfan da parte delle aerolinee, già poco disposte a sperimentare una innovazione tecnica, non esente a priori da rischio.

Fissata quindi la scelta su un motore turbofan, la principale attività nel campo della propulsione durante la progettazione preliminare è la determinazione in collaborazione con i motoristi dei parametri principali del ciclo motore. A questo proposito occorre tenere presente il notevole progresso della tecnica motoristica negli ultimi 10-15 anni. Tale progresso, dovuto ad una molteplicità di fattori, può considerarsi rappresentato soprattutto dal costante e progressivo aumento del rapporto di compressione, che sta passando dal valore tipico di 15:1 della maggior parte dei motori attualmente in servizio, a 20-25:1 dei motori dell'ultima generazione, a circa 35:1 per quelli della prossima. Ciò modifica il punto di ottimizzazione dei parametri del ciclo, in particolare spostato verso valori più elevati il valore ottimo del rapporto di diluizione (Fig. 14). Si sottolinea il fatto che anche in questo caso l'ottimizzazione viene fatta rispetto al DOC sul particolare velivolo. Su un altro tipo di velivolo

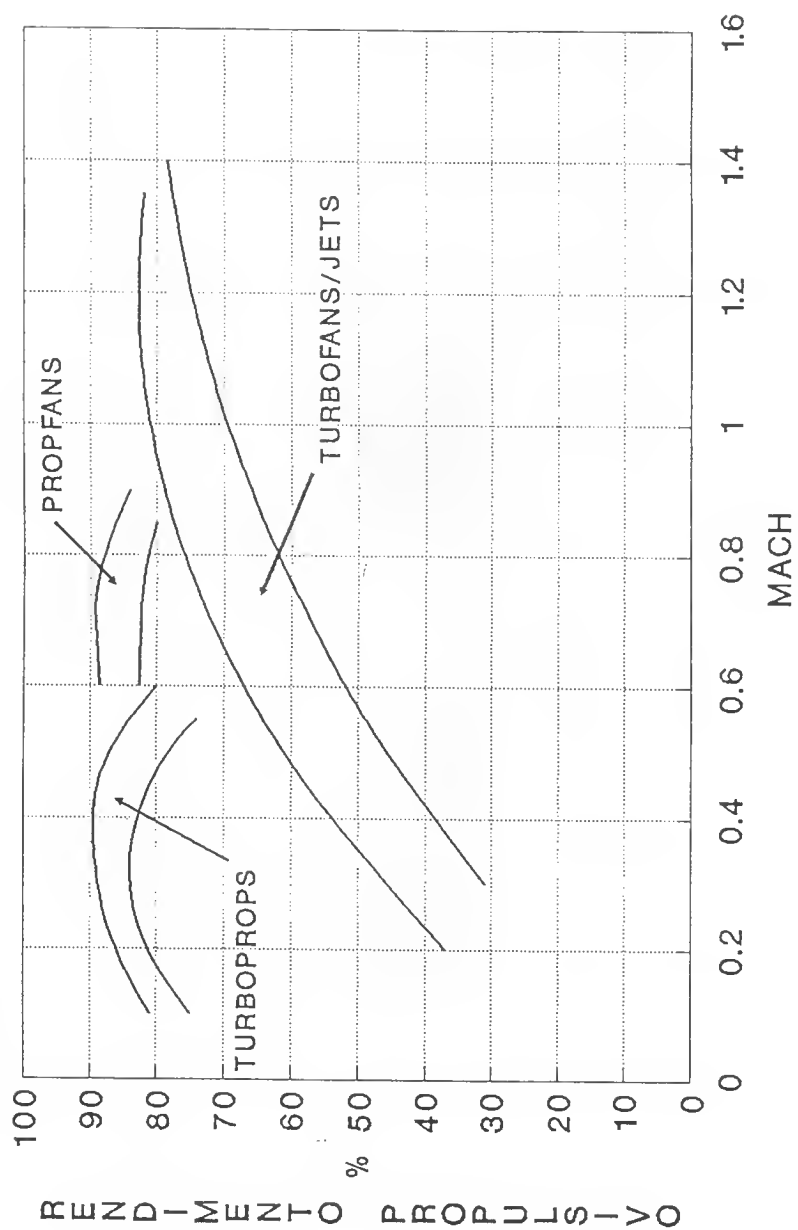


Figura 8 - Confronto rendimento teorico di alcuni tipi di propulsori.

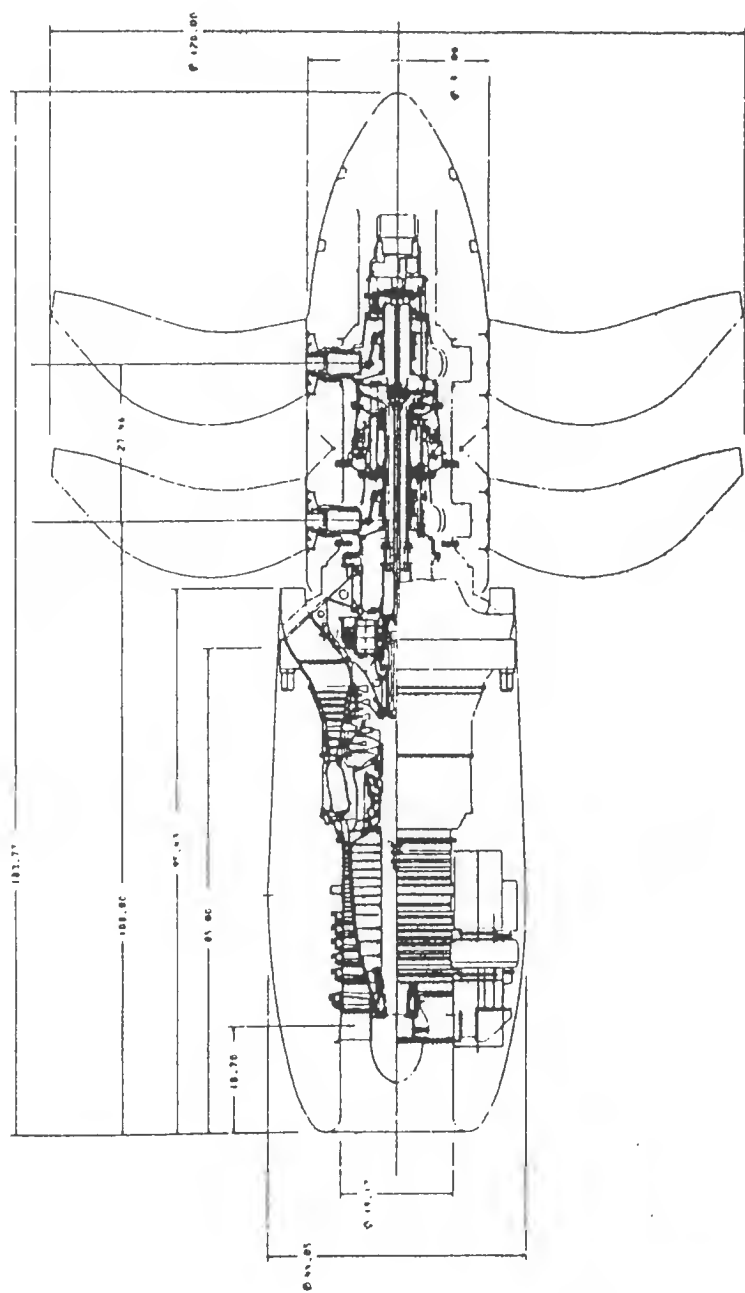


Figura 9 - Propulsore Propfan.

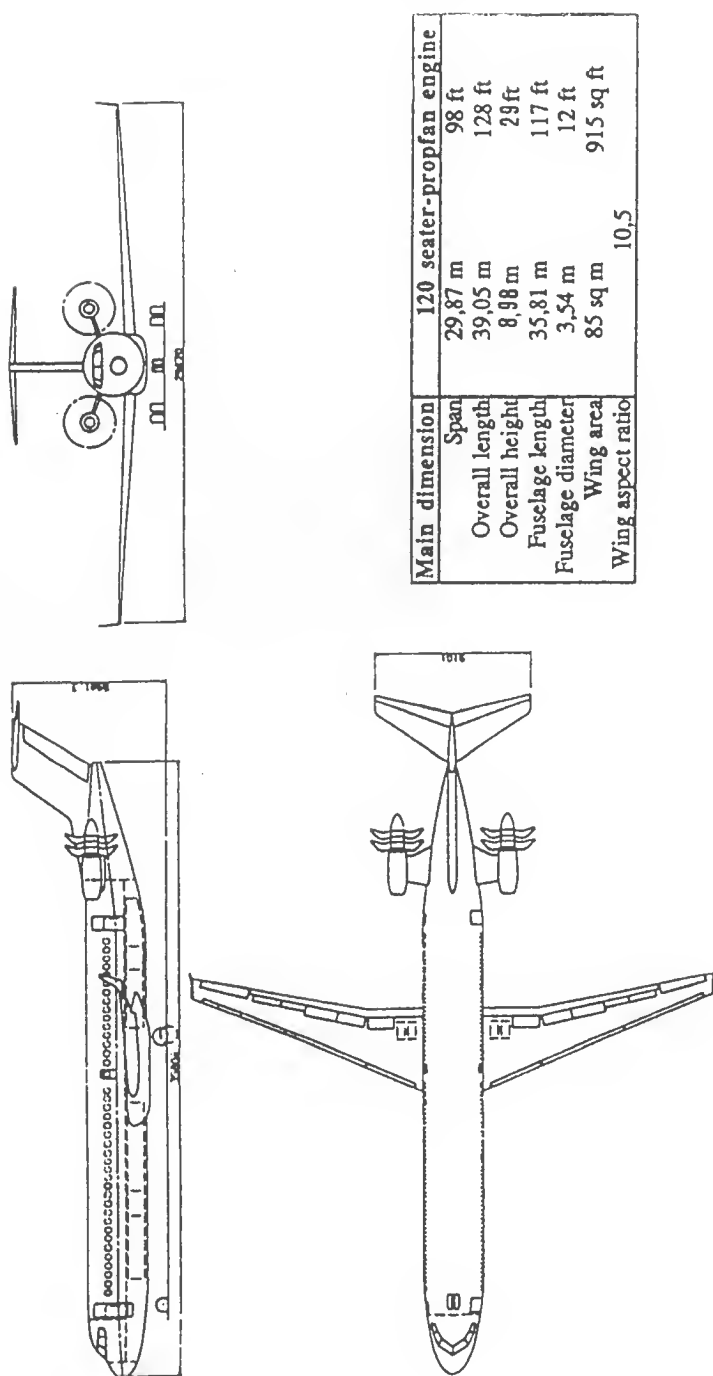


Figura 10 - «120 Seater» - Propfan aircraft.

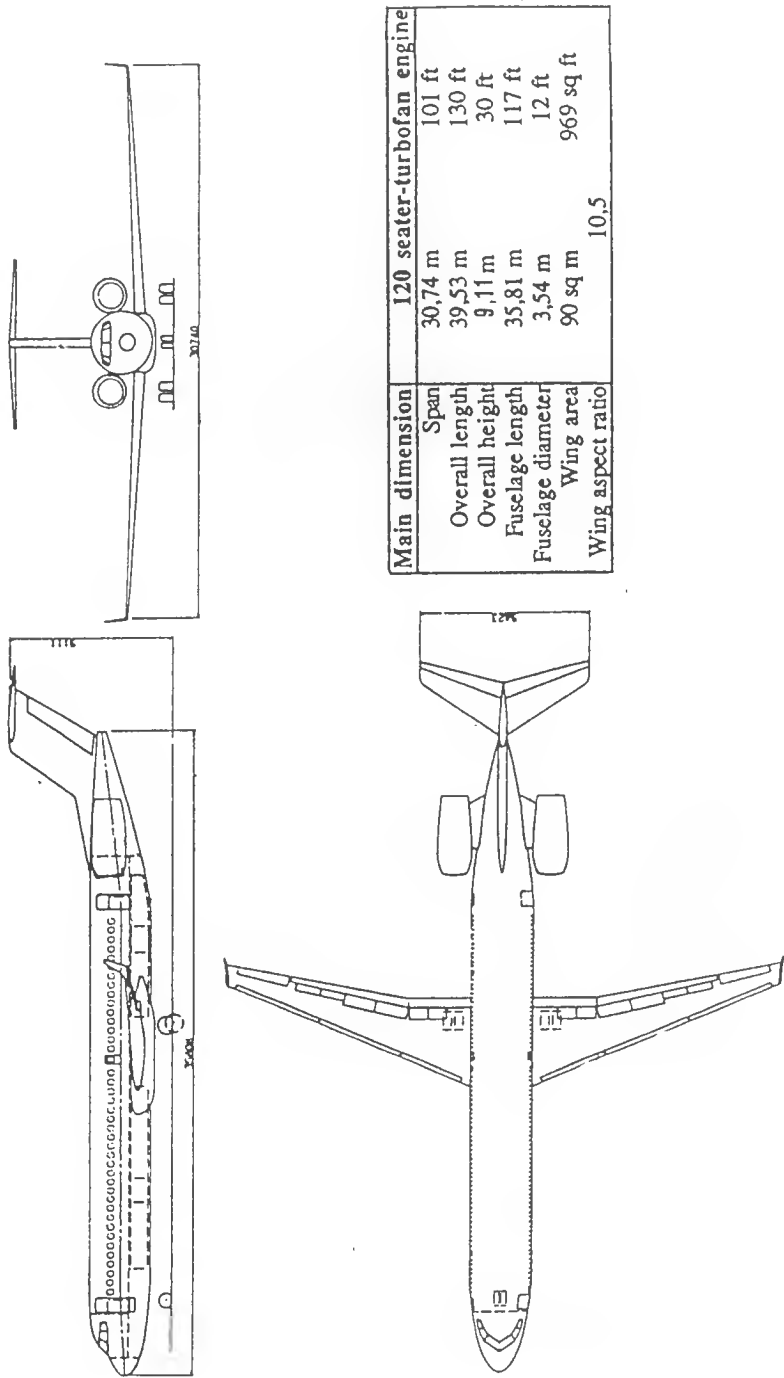


Figura 11 - «120 Seater» - Turbofan aircraft.

400 NM STAGE LENGTH
FUEL 0.75 US\$/Gal
JANUARY 1990 E.C.

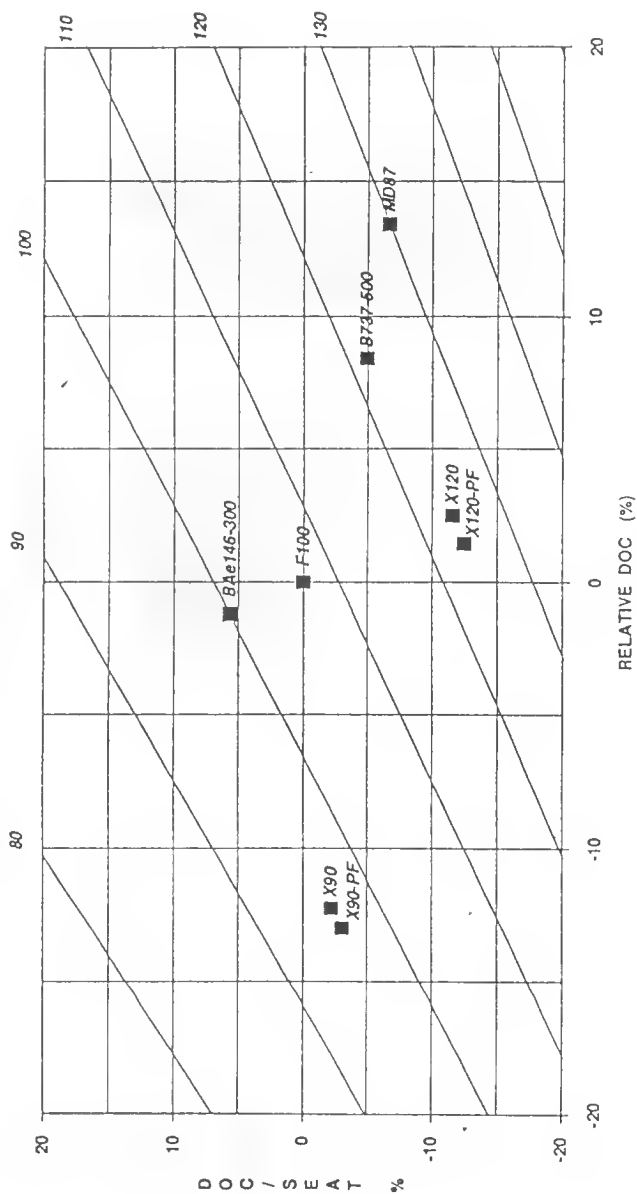


Figura 12 - Direct Operating Cost Comparison.

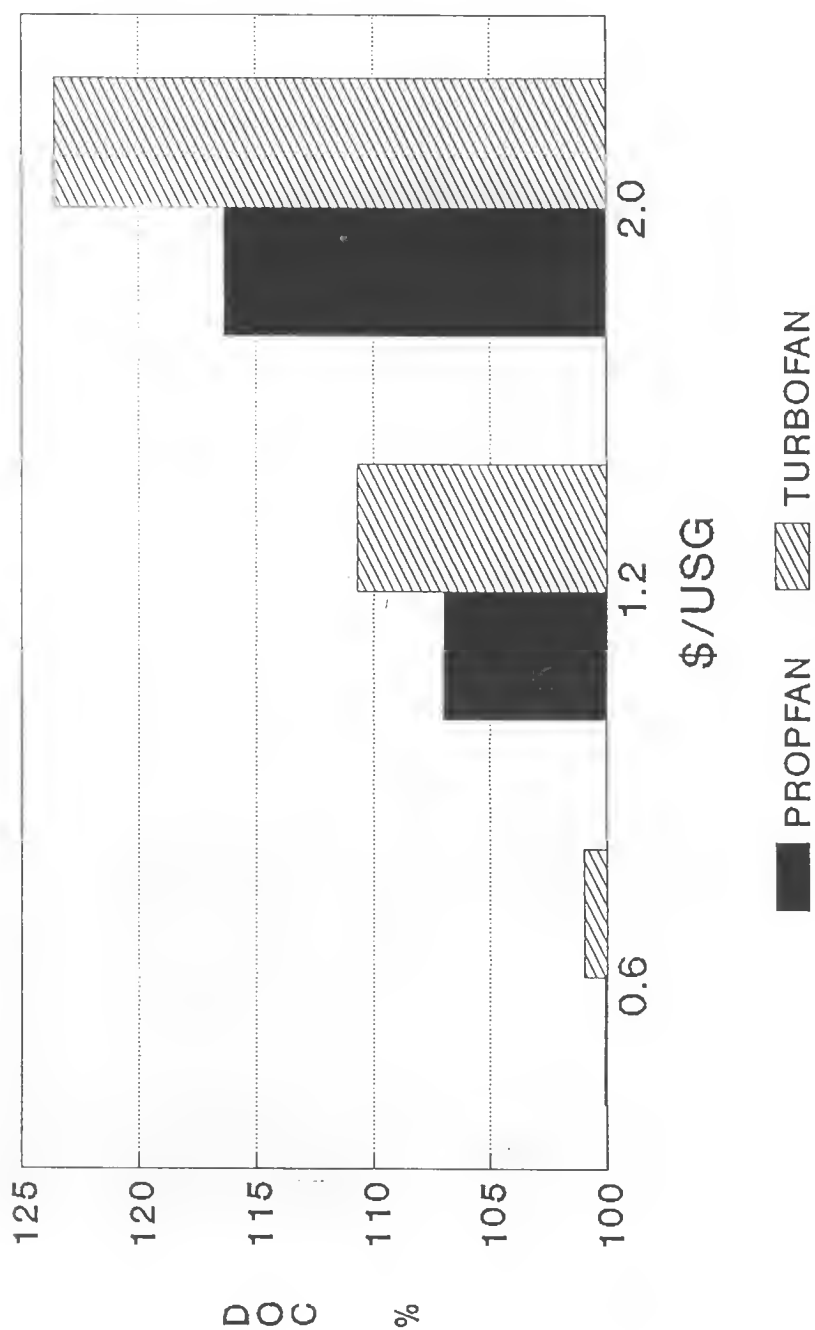


Figura 13 - Effetto sul DOC del prezzo combustibile (tratta tipica = 400 NMI - 120 PAX).

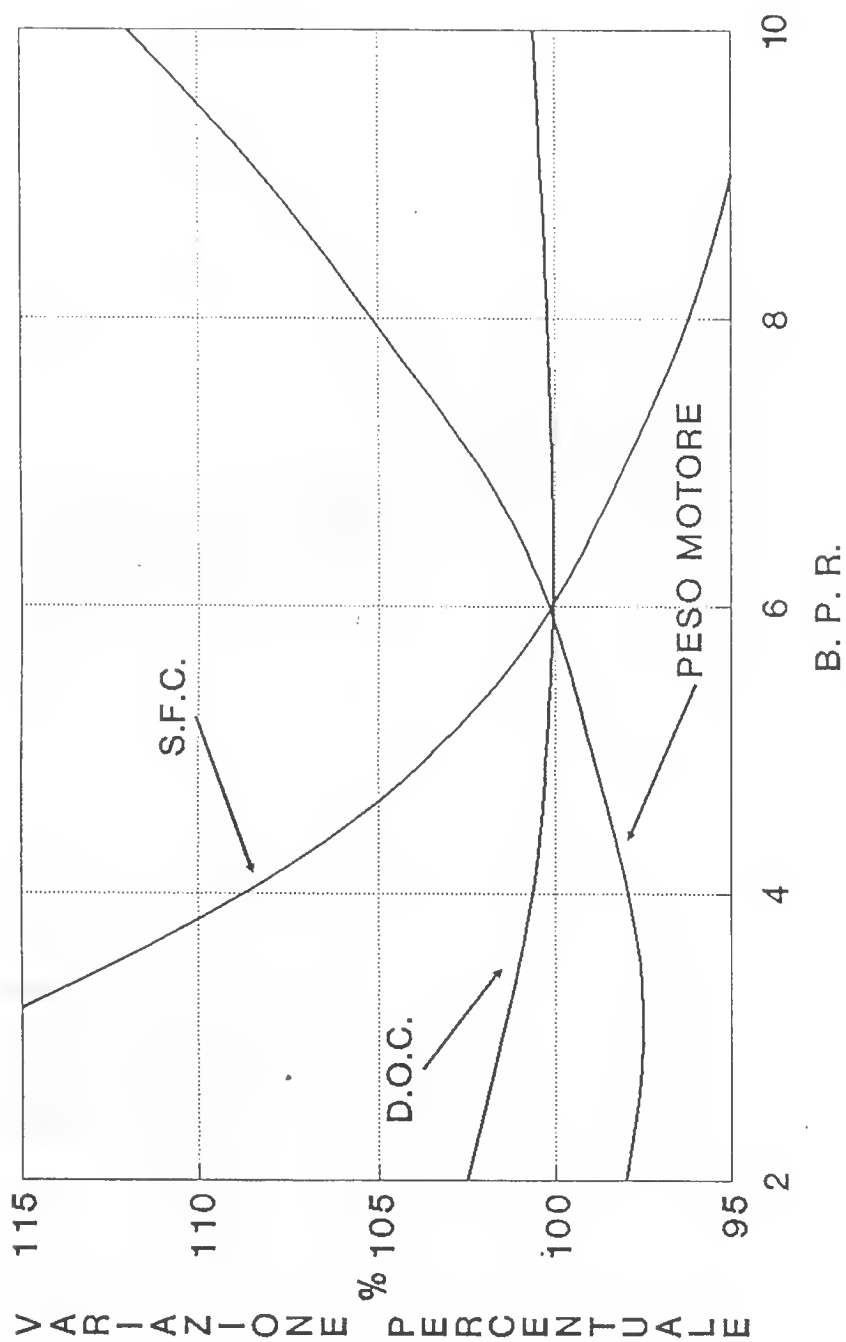


Figura 14 - Effetto del rapporto di diluizione (B.P.R.) del motore.

il risultato sarebbe stato diverso. Ad esempio, un velivolo a lungo raggio avrebbe avuto un valore più elevato del rapporto di diluizione ottimale.

Posizione dei motori. Come noto esistono due soluzioni possibili per l'installazione dei motori in un velivolo bimotore: installazione subalare (B 737, A 300, B 767, A 320, etc.) e installazione lateralmente alla fusoliera posteriore (DC 9, MD 80, BAC 1-11, F 100, etc.).

Non essendo a priori evidente quale fosse la migliore, si è effettuato un approfondito studio di confronto di queste due soluzioni. La Fig. 15 riporta i relativi vantaggi e svantaggi, soprattutto dal punto di vista dell'utente. Sono stati esaminati anche aspetti operativi di dettaglio, ad esempio le Fig. 16 e 17 riportano la situazione del servizio a terra per il 90 posti, aspetto molto importante per ridurre al minimo il tempo di sosta a terra e che ha imposto di aumentare la lunghezza di fusoliera di quasi un metro per la configurazione con motori posteriori.

In termini di peso e di resistenza aerodinamica le due soluzioni sono praticamente equivalenti (Fig. 18 e 19), per cui si può concludere che dal punto di vista tecnico ed economico non vi sono particolari motivi per scegliere una rispetto all'altra. La scelta finale è tuttavia caduta sulla soluzione subalare per la preferenza manifestata per tale soluzione da alcune importanti aerolinee. Tale preferenza è stata motivata con dati quantitativi desunti dall'esperienza sui costi di manutenzione e sui tempi per effettuare interventi sul campo.

Sezione di fusoliera. Si tratta di un elemento molto importante, in quanto influenza direttamente la comodità dei passeggeri, la sistemazione del bagaglio a mano, la facilità di accesso e di uscita e il servizio di bordo. Questo punto è stato lungamente studiato con alcune aerolinee, arrivando infine a fissare le dimensioni indicate in Fig. 20 e 21, dimensioni che, come si vede permettono di realizzare un livello di comodità migliore anche di velivoli più grandi. Tali dimensioni possono essere ottenute sia con una fusoliera circolare che con una a doppio lobo (Fig. 22); la scelta fra queste due sarà fatta in base ai risultati in uno studio attualmente in corso sui costi di produzione. Ovviamente la comodità si paga, come evidenziato dalla Fig. 23.

Configurazione alare. L'ala (Fig. 24) rappresenta l'elemento più importante della cellula e il progetto dell'ala influenza direttamente il dimensionamento generale del velivolo e del motore. Per questo motivo l'ottimizzazione dell'ala deve avvenire già durante la progettazione preliminare. Si tratta di una delle attività classiche dell'ingegneria aeronautica

1. CONFIGURAZIONE

MOTORI SUBALARI (PRO)

- | | | |
|--|---|--|
| • MIGLIORE ACCESSIBILITA' DEL MOTORE PER LA MANUTENZIONE | ➡ | MINORI COSTI DI MANUTENZIONE |
| • MINORE POSSIBILITA' DI DANNEGGIAMENTO DEL MOTORE PER INGESTIONE DI CORPI ESTRANEI (F.O.D.) | | |
| • NON E' NECESSARIO UN PIANO DI CODA A "T". | ➡ | NON CI SONO PROBLEMI DI "SUPER STALLO" |

Figura 15 - Motori in coda Vs motori subalari (Pro & Contro).

1. CONFIGURAZIONE

MOTORI SUBALARI (CONTRO)

* OPERAZIONI DI CARICAMENTO
RESE PIU' DIFFICILI DALLA
MAGGIORE ALTEZZA DA TERRA
DELLA PORTA CARGO

* PIU' ELEVATI LIVELLI DI
RUMORE NELLA FUSOLIERA

MAGGIORI COSTI DELLE
OPERAZIONI A TERRA

NECESSITA' DI MAGGIO
RI TRATTAMENTI ACU -
STICI

2. AERODINAMICA

MOTORI SUBALARI (PRO)

* MIGLIORI CARATTERISTICHE
DI STABILITA' E CONTROLLO

Figura 15 (cont.) - Motori in coda Vs motori subalari (Pro & Contro).

2. AERODINAMICA

MOTORI SUBALARI (PRO)

- * MINORE RESISTENZA DI TRIM-MAGGIO DOVUTA ALLA MINORE ESCURSIONE DEL BARICENTRO

MOTORI SUBALARI (CONTRO)

- * MAGGIORE RESISTENZA IN CRO-CIERA DOVUTA ALLA INTERFERENZA GONDOLE MOTORE/ALA

3. PRESTAZIONI

MOTORI SUBALARI (CONTRO)

- * MAGGIORE SPINTA E' RICHIESTA PER SODDISFARE I REQUISITI IN CONDIZIONI "O.E.I."

Figura 15 (cont.) - Motori in coda Vs motori subalari (Pro & Contro).

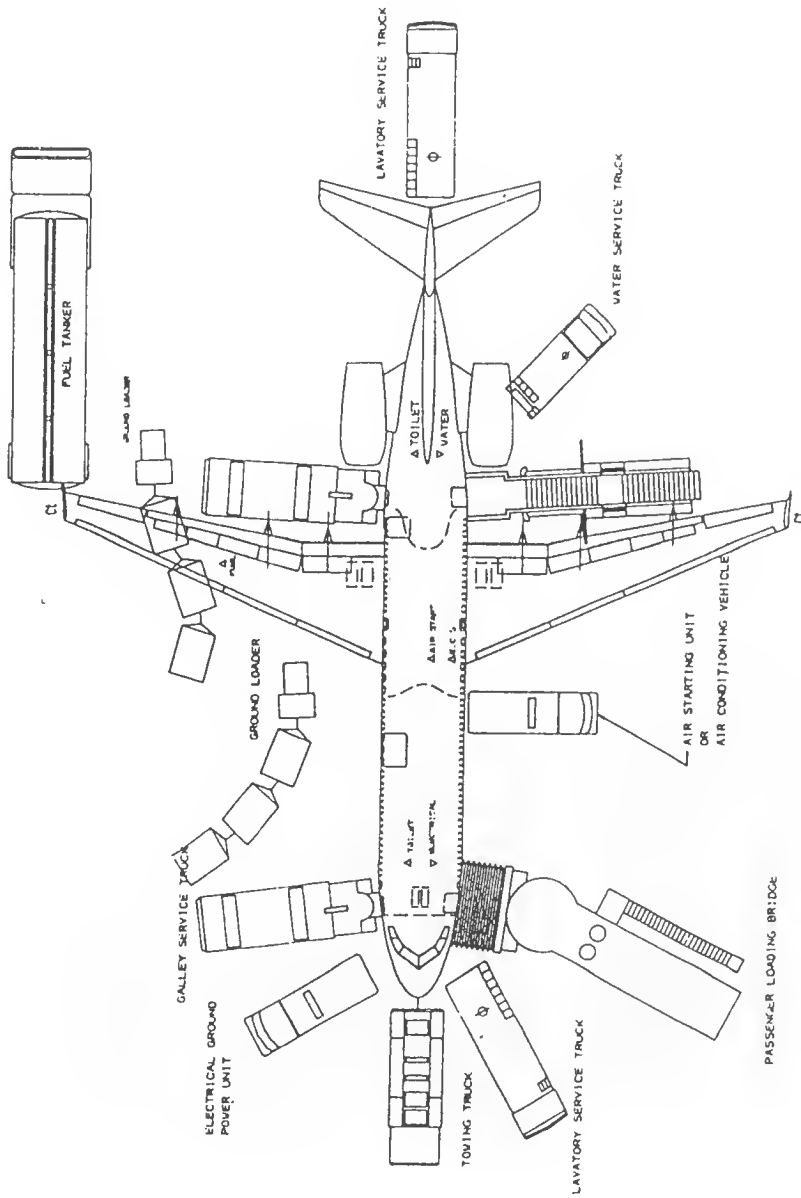


Figura 16

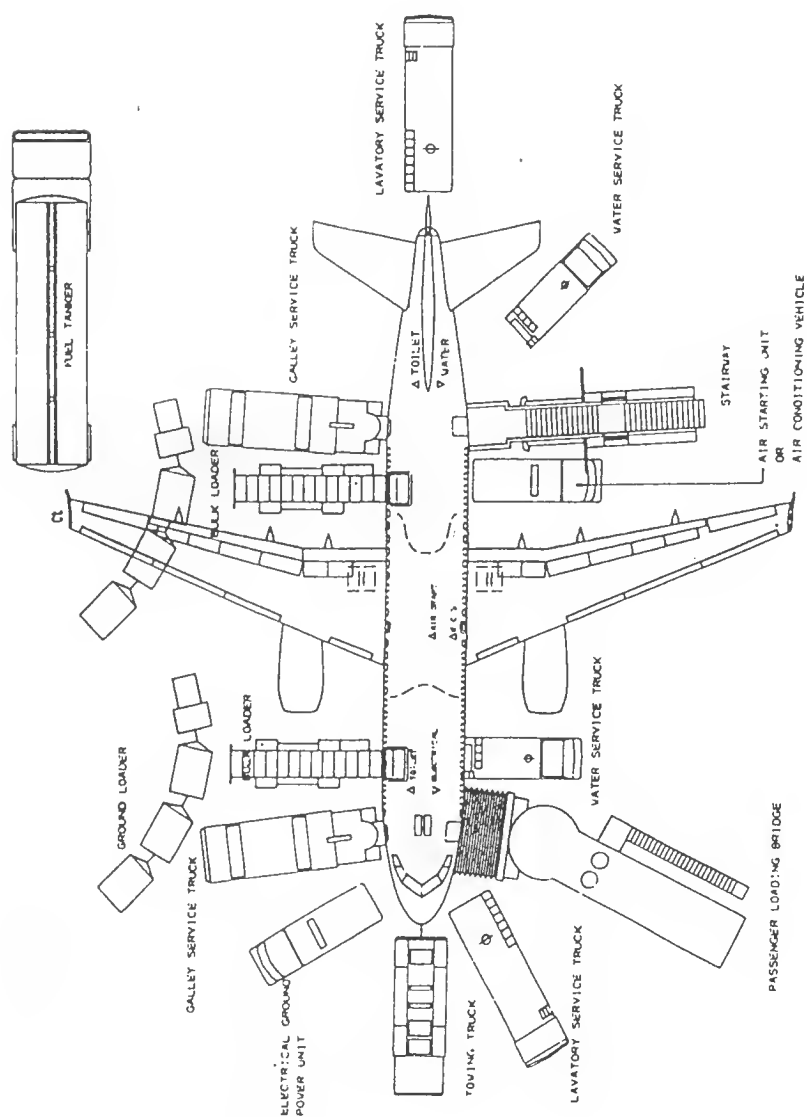


Figura 17

DELTA PESO	
(MOTORE IN ALA - MOTORE IN FUSOLIERA)	
ALA	0 Kg
FUSOLIERA (DELTA LUNGHEZZA = 1M)	- 300 Kg
IMPENNAGGI	- 100 Kg
GRUPPO PROPULSIVO (Motori,piloni,gondole)	+ 250 Kg
CARRELLO	+ 200 Kg
SISTEMI	- 20 Kg

TOTALE	- 30 Kg

Figura 18 - Delta peso.

DELTA CD IN COUNTS

(MOTORE IN ALA - MOTORE IN FUSOLIERA)

+ 5

ALA e GONDOLE

- 2

FUSOLIERA (DELTA LUNGHEZZA = 1M)

+ 3

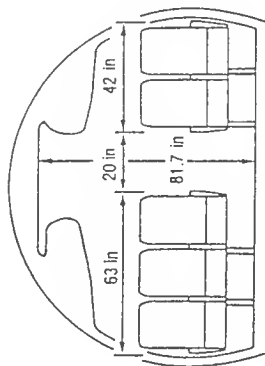
TOTALE

NOTA: RESISTENZA MEDIA TOTALE DEL VELIVOLO = 300 COUNTS

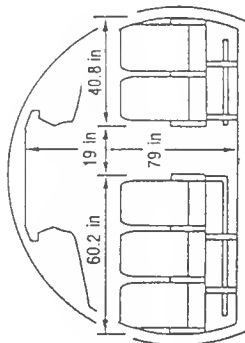
Figura 19 - Delta resistenza.

X-100/200 offre:

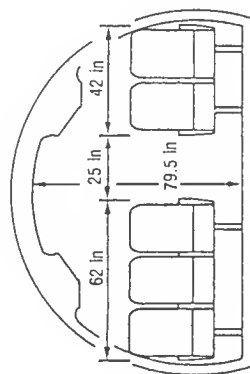
- La migliore altezza in cabina confrontata con i concorrenti a 6 "abreast".
- Ogni sedile e' almeno di un inch plu' largo di quelli del B-737
- "Overhead bins" plu' capienti
- Cabina plu' larga di 5" rispetto alla serie MD80 e di 8" rispetto al F100



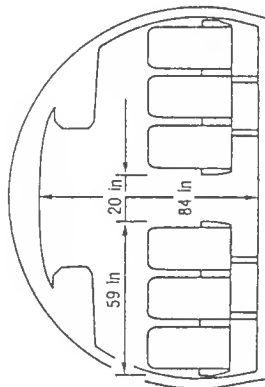
X-100/200



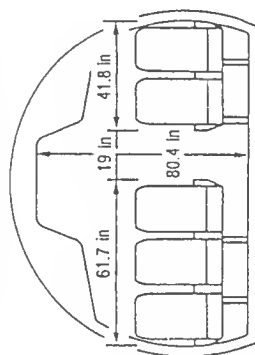
Fokker 100



BAe 146



737 series



MD-80 series

Figura 20 - I vantaggi di una cross section completamente nuova. Cabina passeggeri.

X-100/200 offre:

- Una migliore altezza di lavoro sotto pavimento rispetto al F100 e alla serie MD-80
- Apertura porte cargo verso l'esterno
- La plu' efficiente forma del bagagliaio confrontata con i competitori a 5 "abreast"
- La plu' grande area di accesso

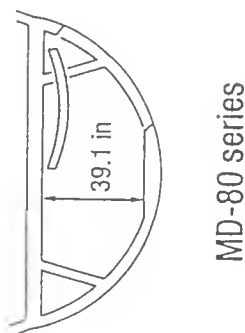
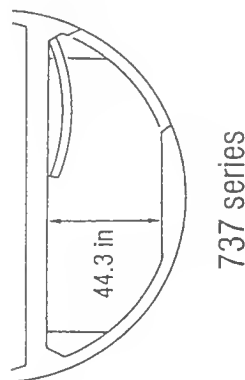
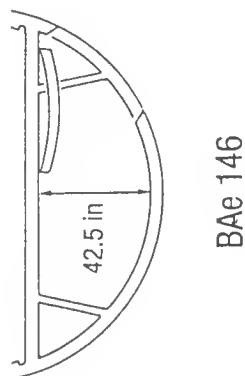
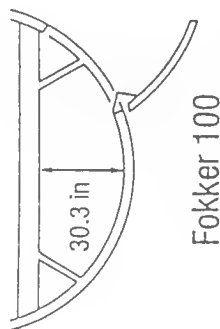
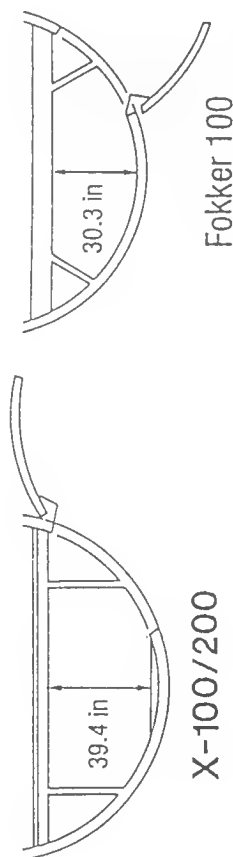


Figura 21 - I vantaggi di una cross section completamente nuova. Vano bagagli.

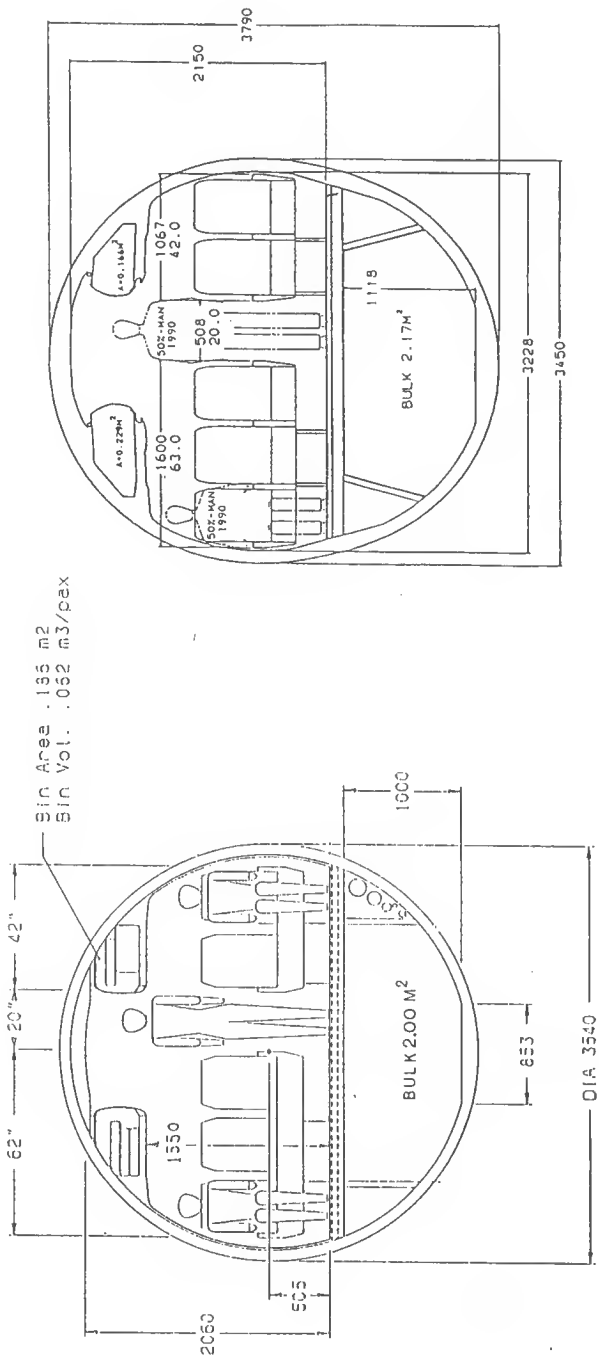


Figura 22 - X-100/200. Cross-section.

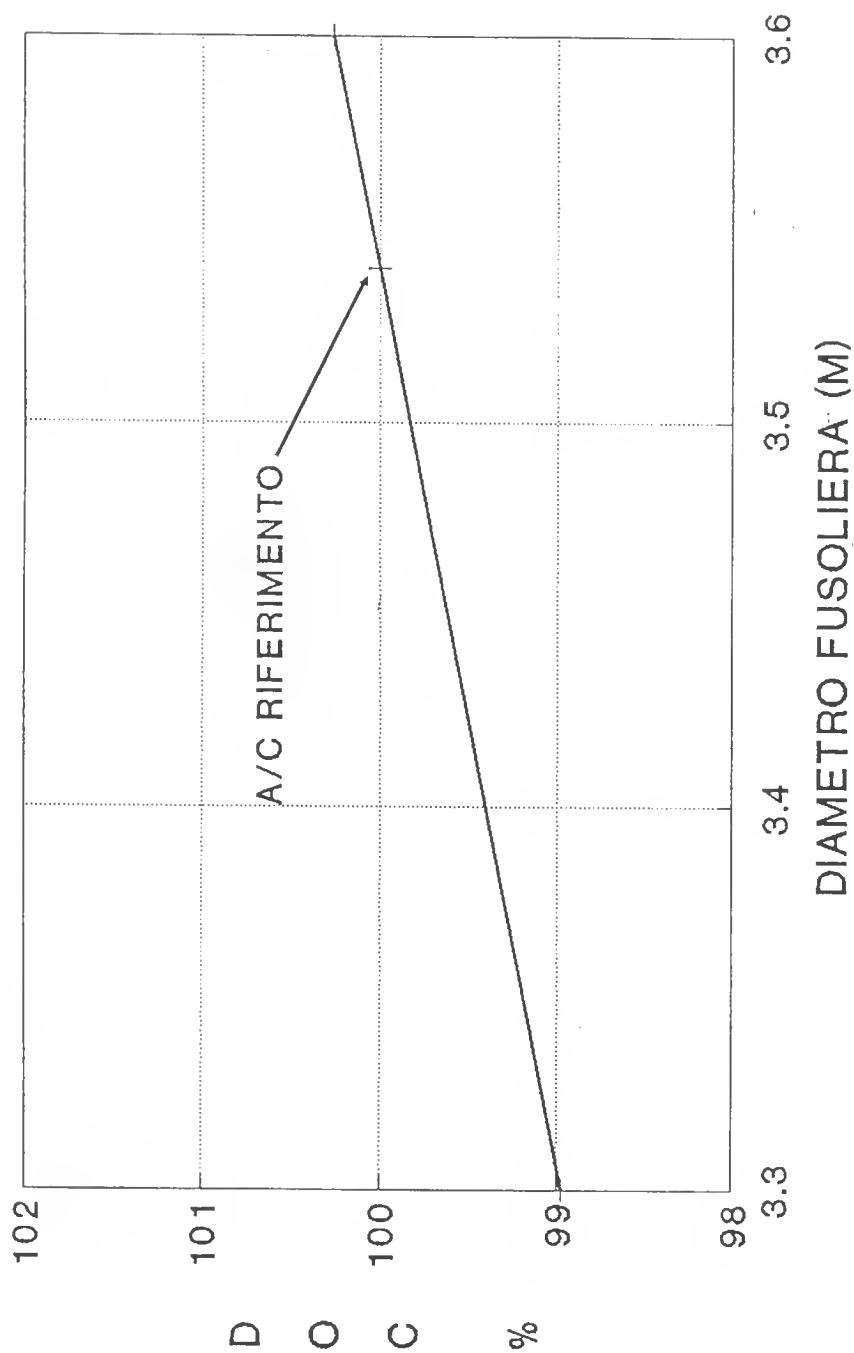


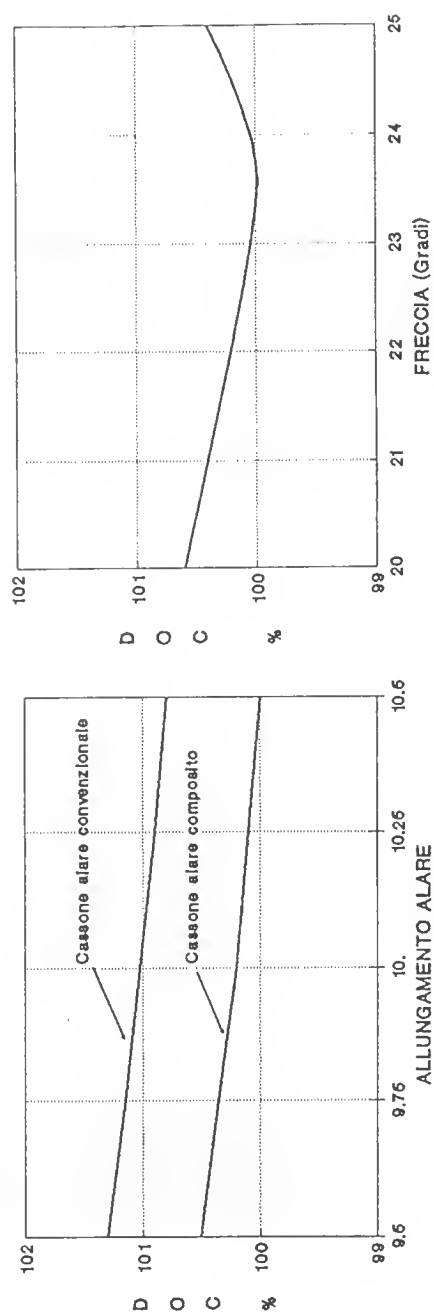
Figura 23 - Effetto sul DOC del diametro fusoliera (tratta tipica = 400 N.Mi - 120 PAX).

e comporta la scelta del profilo (in base a studi precedenti), dell'angolo di freccia, dell'allungamento, della distribuzione dello spessore e dello svergolamento lungo l'apertura e la definizione del sistema di ipersostentazione e delle superfici di controllo. Particolare attenzione deve essere prestata ad ottenere una distribuzione di carico ottimale lungo l'apertura ed una corretta distribuzione di pressione sul dorso dell'ala, per ottenere velocità ed efficienza aerodinamica in crociera. Ovviamente è anche essenziale il raggiungimento degli obiettivi di prestazioni al decollo e all'atterraggio (come sopra accennato, per questo velivolo la superficie alare è determinata dall'atterraggio).

A titolo di esempio in Fig. 25 è riportato l'effetto sul D.O.C. dell'angolo di freccia e dell'allungamento.

Materiali. Anche se si tratta di un argomento non strettamente di configurazione, attualmente la scelta dei materiali per la struttura primaria deve essere fatta già in fase di progettazione preliminare. Infatti essendoci oggi la possibilità di ricorrere ai materiali compositi, che consentono una riduzione apprezzabile del peso a vuoto del velivolo, la scelta dei materiali influenza direttamente il dimensionamento generale del velivolo. Ovviamente il vantaggio della riduzione di peso va valutato a fronte delle conseguenze sui costi di produzione. I materiali compositi sono notevolmente più costosi di quelli metallici tradizionali, per cui bisogna concepire progetto e processo produttivo in modo da recuperare nella fabbricazione e, soprattutto, nell'assemblaggio delle strutture. È quindi evidente che una decisione sull'applicazione dei materiali compositi ad un elemento strutturale può essere presa solo dopo uno studio di un certo dettaglio che comporta la definizione di massima del progetto, del processo produttivo e dei costi, da confrontare ad una soluzione metallica tradizionale. Per quanto riguarda il rischio tecnico, invece, la tecnologia delle strutture in composito è già sufficientemente matura da consentire l'applicazione a strutture primarie senza particolari rischi nello sviluppo, nella certificazione ed in servizio.

La Fig. 25 mostra il vantaggio in termini di DOC di un cassone alare in composito. Come si vede, esso non è molto importante, per cui bisogna fare il massimo sforzo nella fase realizzativa per migliorare la situazione in termini di costi di produzione. La Fig. 26, infine, riporta il quadro completo dell'utilizzazione di materiali compositi prevista per questo velivolo.



BFL - 5000 Ft
MMO - .81

Figura 25 - Effetto sul DOC dell'allungamento e della freccia (tratta tipica = 400 N.Mi - 120 PAX).

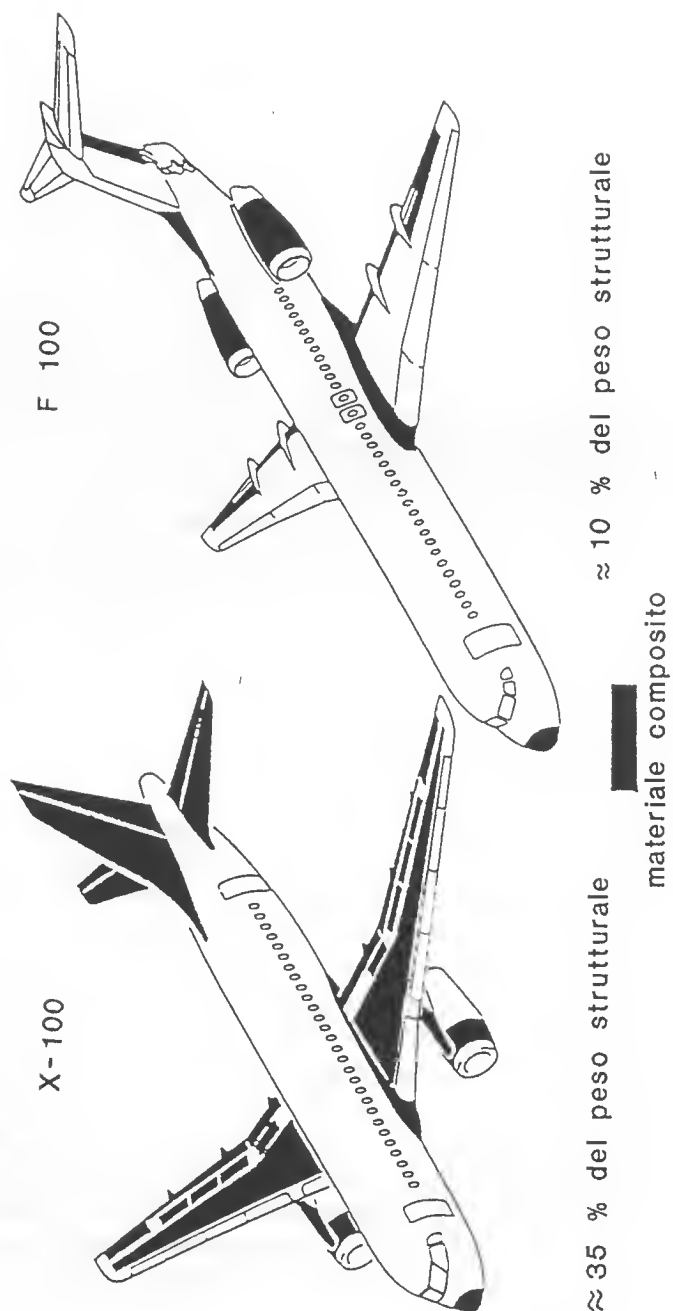


Figura 26 - Utilizzazione materiali compositi.

Conclusioni

Come risultato degli studi di ottimizzazione di cui in precedenza si è cercato di dare qualche esempio, si è pervenuto a finalizzare la configurazione dei due velivoli della famiglia iniziale (Fig. da 27 a 31). Certamente, a nessuno sarà sfuggita l'entità molto limitata del guadagno in termini di D.O.C. apportato dall'ottimizzazione di ogni singolo parametro di progetto e di ogni singola tecnologia.

Da questa constatazione deriva un primo aspetto che, come conclusione, si vuole mettere in rilievo: se si vuole ottenere un risultato significativo in termini di riduzione del D.O.C., è necessario che il lavoro di ottimizzazione venga applicato ad un numero grandissimo di argomenti. In altre parole, questo lavoro deve essere paziente, tenace e puntiglioso nell'andare a considerare ogni possibile aspetto del progetto, che possa dare un ritorno. L'opera, inoltre, deve essere proseguita, oltre la fase di progettazione preliminare, nella definizione finale, nella progettazione esecutiva e nella messa a punto del prodotto, andando a interessare anche i minimi dettagli.

A proposito di tali fasi successive del progetto, si ritiene opportuno accennare alle linee direttrici, secondo le quali deve svolgersi in esse il lavoro di ottimizzazione tecnico-economica del prodotto.

Per questo è necessario riferirsi di nuovo alla Fig. 3, dalla quale risulta che la componente più importante del D.O.C. è rappresentata dagli oneri finanziari, costituiti in gran parte dell'ammortamento del costo del velivolo. Tale costo è costituito ovviamente dalle seguenti componenti principali (in ordine di importanza):

- costi di produzione (costi interni di trasformazione e costi di acquisto per materiali ed equipaggiamenti),
- ammortamento dell'investimento iniziale (progettazione, attrezzamento, certificazione, etc.),
- costi "ricorrenti" di gestione (commerciali, per l'assistenza post-vendita, etc.).

Poiché per la seconda componente un velivolo di nuovo progetto risulta in genere svantaggiato (soprattutto per le esigenze sempre crescenti nel tempo per sicurezza, certificazione e, in generale, qualità del prodotto), appare evidente l'importanza fondamentale per un nuovo velivolo di ridurre i costi di produzione. Se non si riesce a dimostrare di potere contenere tali costi (e quindi il prezzo del velivolo) entro limiti stabiliti, il programma di sviluppo di un nuovo velivolo commerciale non risulta fattibile dal punto di vista economico.

Per ottenere questo risultato sono oggi disponibili dei mezzi che non

esistevano quando furono sviluppati i velivoli della generazione precedente e che possono risultare in grandi benefici in termini di riduzione dei costi. A tal fine, si deve operare secondo tre linee d'azione principali:

- A - applicare le tecniche di ottimizzazione sopradescritte anche al sistema propulsivo, ai sistemi e agli equipaggiamenti,
- B - progettare contemporaneamente il *prodotto*, il *processo produttivo* e il *processo manutentivo*,
- C - applicare largamente al processo industriale metodi di automazione e di gestione computerizzata.

Circa il punto A, si rammenta che quelli che abbiamo chiamato costi di produzione in realtà si compongono per circa il 50% di costi di acquisto per i motori e gli equipaggiamenti. Il costo dei motori da solo rappresenta circa il 20% del costo totale del velivolo. Di qui la necessità di incidere anche su queste voci.

L'elemento comune, che rende possibili operare secondo le tre linee sopracitate, e la disponibilità delle moderne tecnologie basate sul calcolatore numerico. Ad esempio, per quanto riguarda il punto B, è la disponibilità dei moderni sistemi CAD/CAM, che consente di rendere disponibili in maniera continua e «in tempo reale» i dati di progetto, prima del loro congelamento, alle altre organizzazioni a valle della progettazione (ingegneria industriale, ingegneria di manutenzione, etc.). Tali organizzazioni possono quindi operare parallelamente alla progettazione, immettendo in essa le loro esigenze. In tal modo il prodotto nasce ottimizzato anche dal punto di vista dei costi di produzione e delle esigenze di manutenzione in servizio. La stessa cosa può dirsi per il punto C: il prodotto deve essere progettato appositamente per potere essere prodotto con determinati metodi automatizzati.

Questi fattori stanno determinando una rivoluzione nella produzione aeronautica, che, partendo da una situazione tradizionale quasi artigianale, verrà sempre più assumendo caratteristiche di complessità e di sofisticazione, finora tipici delle produzioni di grandi serie.

Infine, tornando alla Fig. 3, si può rilevare che la seconda componente in ordine di importanza del D.O.C. (a parte l'equipaggio) è rappresentata dai costi di manutenzione. Di qui l'importanza di sviluppare un progetto che renda minimi tali costi: la facilità e il basso costo di manutenzione debbono essere uno dei principali requisiti di progetto.

In alcuni casi, la manutenzione può determinare alcune scelte fondamentali di progetto, come nel caso precedentemente discusso del posizionamento dei motori.

Nei limiti imposti a questa esposizione, si è cercato di dare un'idea

di massima sui criteri, che si debbono seguire nella concezione di un velivolo da trasporto commerciale, per ottenere un prodotto di successo.

Per riepilogare, tali criteri possono sinteticamente riassumersi nei seguenti punti principali:

- 1) Si deve operare in modo integrato sui requisiti delle aerolinee, sulle tecnologie e sul progetto del velivolo.
- 2) Il grado di ottimizzazione del prodotto per qualità ed economia deve essere estremamente spinto.
- 3) Il costruttore del velivolo, i costruttori del motore e degli equipaggiamenti principali e, soprattutto, le aerolinee future utenti del velivolo debbono lavorare insieme nella concezione e nello sviluppo della macchina.

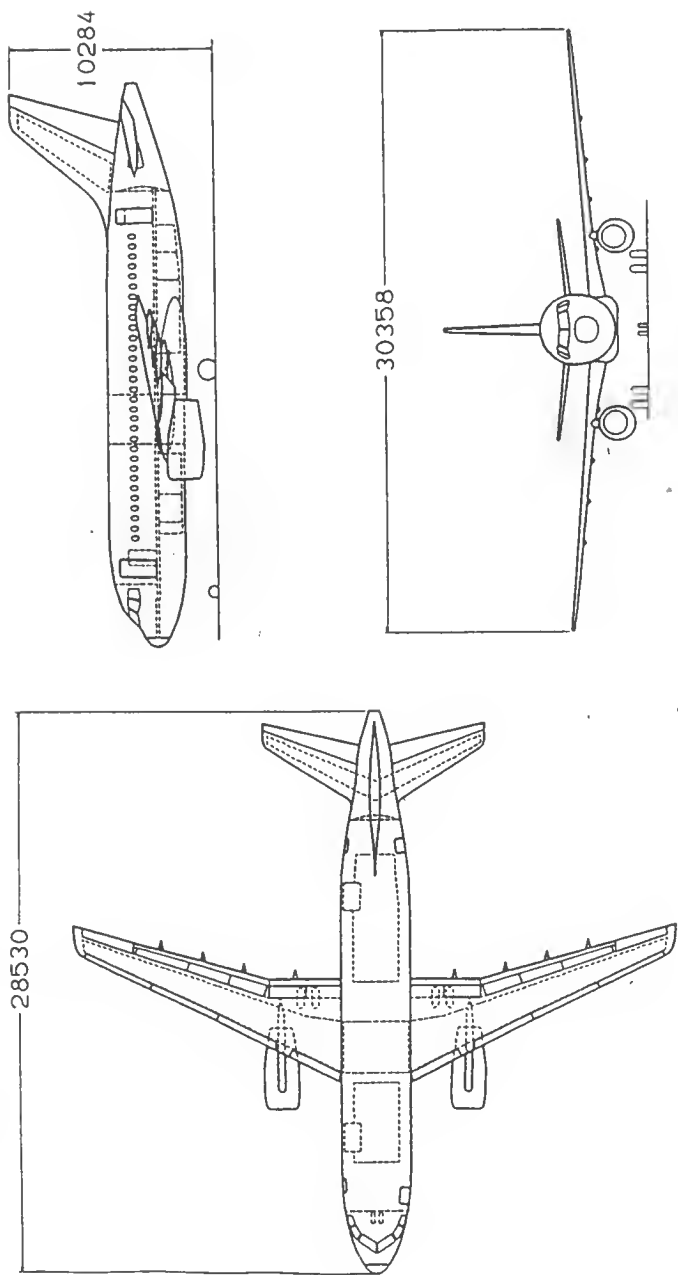


Figura 27 - X-100 General arrangement.

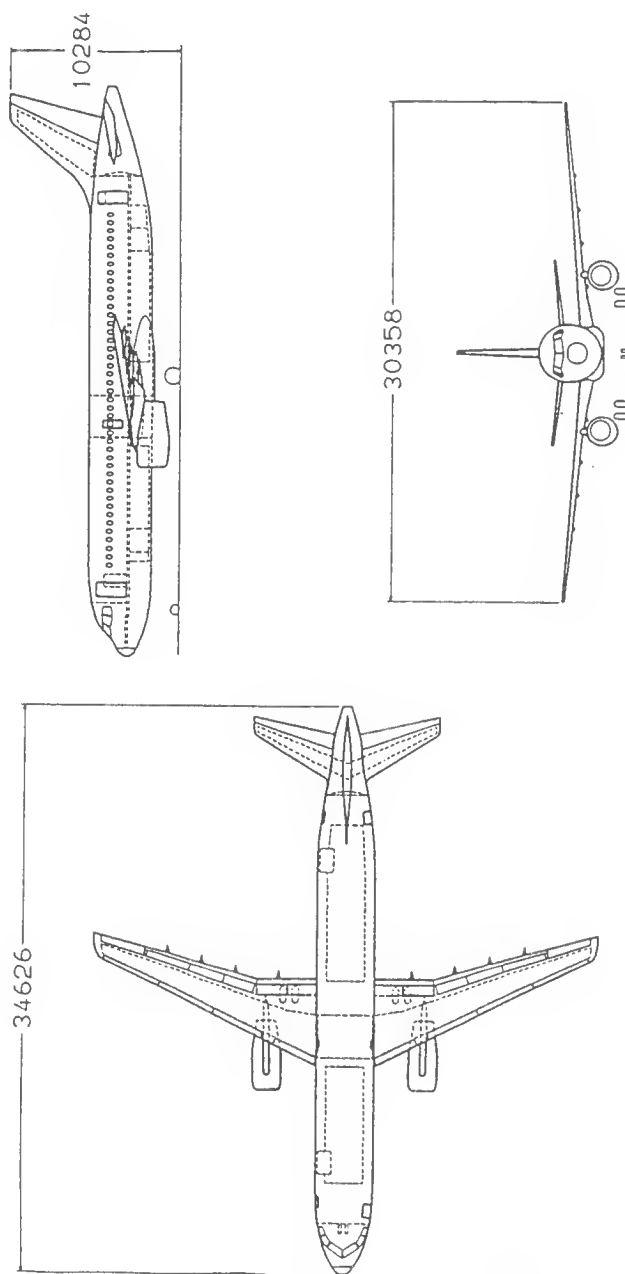


Figura 28 - X-200 General arrangement.

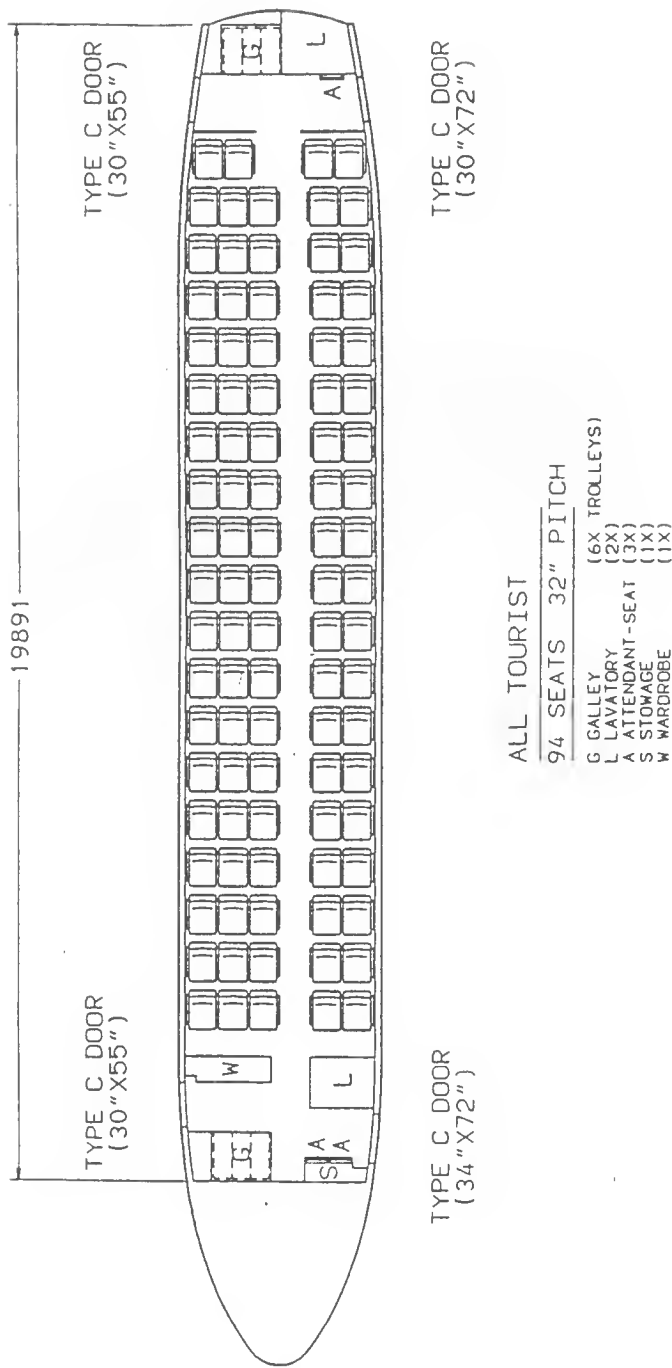


Figura 29 - X-100 Cabin layout.

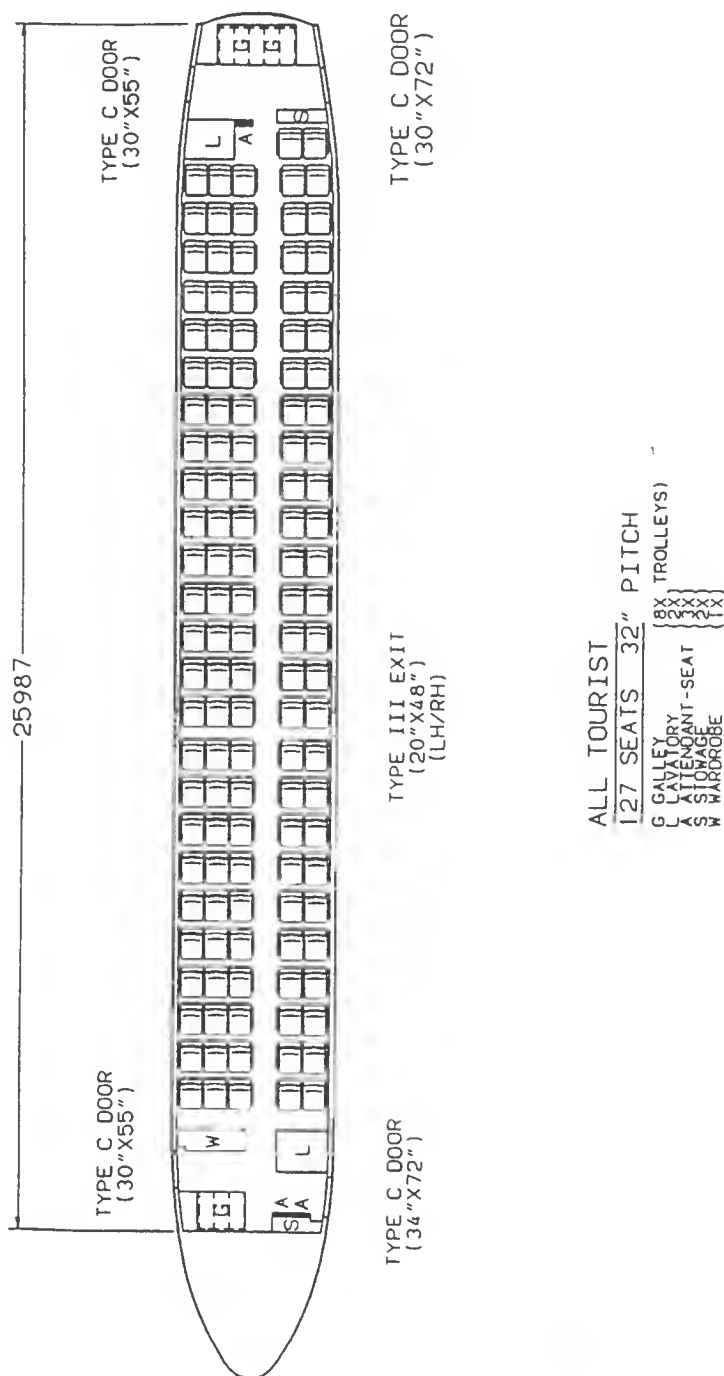


Figura 30 - X-200 Cabin layout.

	<u>X-100</u>	<u>X-200</u>
- PESO MASSIMO, KG	43.000	49.800
- CARICO PAGANTE MASSIMO, KG	10.200	14.800
- SPINTA MOTORE, lb	15.500	18.500
- DIST. DI DECOLLO BILANCIATA (*), ft	4.400	5.000
- DISTANZA DI ATTERRAGGIO (*), ft	3.900	4.400
- VELOCITA' DI AVVICINAMENTO, kts	116	125
- VELOCITA' MASS. DI CROCIERA, Mach	0.81	0.81
- AUTONOMIA A MASS. CARICO mn	1.500	1.500
- RUMORE	25EPNdB sotto FAR25	part 36 stage III

(*) Condizioni standard, l.m.

Figura 31 - Caratteristiche principali.



Sulla dinamica del volo controllato

Attilio SALVETTI, Giovanni MENGALI

Introduzione

La dinamica del volo del velivolo controllato, nella sua attuale formulazione, è il risultato di un lungo e complesso processo in cui le capacità progettuali e costruttive, gli sviluppi teorici nel campo della meccanica, dell'aerodinamica e dei controlli ed i metodi di sperimentazione in volo, hanno gradualmente prodotto un complesso integrato di conoscenze che costituisce il nucleo centrale delle metodologie avanzate oggi in uso nel campo della progettazione aeronautica.

Il prof. Gabrielli è stato un uomo che, certamente, anche con molto anticipo rispetto ai tempi, ha contribuito a questo fondamentale processo di sintesi, riassumendolo nella sua persona e trasferendolo nelle sue significative realizzazioni in campo aeronautico.

Per riassumere il processo secondo il quale il complesso delle conoscenze nelle diverse aree si è andato evolvendo, vengono innanzi tutto ricordate alcune delle tappe più significative della prima fase di questo sviluppo quale si è venuta configurando nei primi 50 anni di evoluzione dell'aeronautica. Le successive fasi, cioè quelle che prendono le mosse a partire dalla fine degli anni quaranta, vengono messe in luce attraverso la presentazione di alcuni risultati sulla caratterizzazione del comportamento dinamico di aeroplani delle ultime generazioni, ottenuti applicando metodologie e tecniche che si sono venute consolidando in quest'ultimo periodo.

1. Primi sviluppi della dinamica del volo controllato

La dinamica del volo, nella sua attuale concezione, ha come oggetto principale d'indagine il comportamento del velivolo controllato. Si tratta di un'area fortemente applicata: uno comprendente lo studio classico sulla stabilità ed il controllo degli aeroplani, un secondo in cui si analizza il comportamento del sistema aeroplano-pilota umano, ed un terzo costituito dalle applicazioni della teoria dei controlli automatici allo sviluppo di sistemi artificiali di stabilizzazione e controllo.

Questo modo di concepire la dinamica del volo si è andato affermando con intensità crescente a partire dalla fine degli anni quaranta in concomitanza con l'avvio di quell'imponente processo di sviluppo tecnologico che da allora ha accompagnato ininterrottamente l'evoluzione dell'ingegneria aeronautica.

Il processo di integrazione, che ha origini antecedenti al primo volo sopra le dune di Kitty Hawk [1], e che ha portato la dinamica del volo a diventare uno degli aspetti più importanti nel progetto degli aeroplani, è stato molto lento ed ha richiesto un primo lungo periodo di gestazione. Durante questo primo periodo le tre aree disciplinari sopra menzionate si svilupparono seguendo strade indipendenti, senza che vi fossero sostanziali comunicazioni fra i loro specialisti.

In particolare le ricerche nel campo della dinamica dell'aeroplano furono prevalentemente indirizzate allo studio della stabilità del volo. Già nel 1904 Bryan e Williams [2], utilizzando l'ipotesi delle piccole perturbazioni, svilupparono le equazioni linearizzate del moto ed introdussero la descrizione delle forze aerodinamiche in termini di derivate aerodinamiche, secondo un approccio ancor oggi seguito. Le indagini sperimentali condotte al NPL da Bairstow e Melvill Jones [3] consentirono fin dal 1912 di disporre di tecniche per la misura delle derivate aerodinamiche, che costituirono la base, per lungo tempo, per i successivi sviluppi sperimentali, al di qua ed al di là dell'Atlantico.

La graduale integrazione dei modelli teorici con i dati sperimentali permise di ottenere attendibili informazioni sul comportamento dinamico del velivolo perturbato, che trovarono anche significative conferme nei risultati delle prime prove di volo in piena scala.

Questa attività di ricerca, sia teorica che sperimentale, cui contribuirono un numero progressivamente crescente di ricercatori sia europei che nordamericani è ben documentata nei volumi di Bairstow [4] del 1920 e di Melvill Jones [5] del 1935. Il contenuto di quest'ultimo volume, di particolare interesse per la sua profondità e modernità, fornisce un quadro chiaro dello stato delle conoscenze del tempo; appare evidente in particolare come l'uso delle equazioni linearizzate del moto si fosse venuto ormai consolidando come un procedimento attendibile per predire la stabilità del moto, per valutare la storia temporale conseguente ad un disturbo o ad un'azione di comando e per analizzare gli effetti di cambiamenti di configurazione sulle caratteristiche di stabilità del volo.

Nei dieci anni successivi, le ricerche in questo campo continuarono con applicazioni ad una gran varietà di aeroplani in diverse condizioni di volo e soprattutto con un significativo progresso nelle analisi delle risposte ai comandi. Più in particolare i risultati che si ottennero riguar-

davano soprattutto:

- il moto del velivolo a comandi bloccati;
- il moto a comandi liberi;
- la risposta a comandi mossi secondo assegnate leggi temporali.

Questi importanti sviluppi nell'analisi dinamica del velivolo non trovarono, tuttavia, un apprezzabile interesse da parte dei progettisti per la difficoltà di associare ai comportamenti dinamici, evidenziati dalle analisi, adeguati criteri di progetto che garantissero la pilotabilità dell'aeroplano.

La fondamentale difficoltà di interpretare sul piano progettuale i risultati relativi alla stabilità del moto dell'aeroplano a comandi liberi o bloccati o in risposta a loro particolari movimenti, nasceva dalla mancanza sul piano teorico di strumenti adatti ad interpretare la funzione del pilota. Ciò che mancava in quegli anni era la capacità di capire che l'aeroplano è solo una parte di un sistema che include il pilota come sua parte essenziale e che la dinamica della sola parte aeroplano può essere molto diversa da quella dell'intero sistema, essendo quest'ultimo profondamente condizionato dalla condotta del pilota.

Sul piano concreto i progettisti erano consci delle capacità del pilota di incidere profondamente sul comportamento dinamico del velivolo. Non era un caso che il primo volo fosse stato condotto con un aeroplano instabile sul piano longitudinale e che molti altri velivoli, realizzati nei vent'anni successivi, potessero essere pilotati pur in presenza di significative forme di instabilità nel moto perturbato a comandi bloccati [6]. Il problema consisteva nella difficoltà di quantificare gli effetti del pilota sul comportamento dinamico del velivolo e di apprezzare quali caratteristiche del moto perturbato fossero significative ai fini della pilotabilità.

La via che si cominciò a seguire in quegli anni fu quella della sperimentazione in volo e della «misura» delle cosiddette qualità di volo, attraverso il confronto fra le caratteristiche di stabilità misurate in volo ed il giudizio dei piloti collaudatori [6], [7]. Questo modo di procedere era caratterizzato da forti limitazioni, dovute alla complessità delle prove di volo, alla difficoltà di una loro razionale concezione a causa della non chiara comprensione del sistema «uomo-macchina» ed alla non sempre soddisfacente professionalità dei collaudatori, tutti piloti eccellenti, ma frequentemente privi dell'indispensabile bagaglio tecnico. Nonostante ciò, il progressivo accumulo di conoscenze culminava, nel 1943, dopo circa vent'anni di difficile sperimentazione, nella pubblicazione dell'ormai classico NACA TR 775 «Requirements for Satisfactory Flying Qualities of Airplanes» cui facevano seguito le prime specifiche militari sulle qualità di volo [8], [9].

Nasceva così il primo corpo di norme per il progetto che sarebbe servito come base per garantire adeguate caratteristiche di pilotabilità ai nuovi aeroplani e che avrebbe costituito il punto di partenza di un nuovo complesso processo di valutazione delle qualità di volo che portò nel 1954 alla pubblicazione della MIL-F-8785 «Military Specification, Flying Qualities of Piloted Airplanes», alla seconda rivoluzionaria versione del 1969, la 8785-B, alla terza revisione, la 8785-C del 1980 ed infine all'ultimo profondo cambiamento della fine degli anni '80 con l'abbandono della MIL-SPEC e l'adozione della MIL-STD con il relativo manuale di supporto [10].

Tornando al NACA TR 775, questo era stato sviluppato prescindendo quasi totalmente dai risultati disponibili sul moto del velivolo perturbato e ciò principalmente a causa dell'assenza di un approccio integrato al problema.

Un altro aspetto della storia è quello legato allo sviluppo dei sistemi di controllo automatici ed in particolare agli autopiloti. In questo campo, in cui le prime realizzazioni sono state addirittura antecedenti il primo volo [1], [11], lo sviluppo dei sistemi di controllo automatici è da attribuirsi alla straordinaria capacità ingegneristica di alcuni uomini tra cui primeggiano le figure degli Sperry negli Stati Uniti, il gruppo di ricercatori attivo in Inghilterra presso il RAE e gli ingegneri tedeschi Siemens ed Askania. Senza voler entrare in alcun modo nel dettaglio di questa affascinante parte della storia del volo controllato, che si trova peraltro descritta in alcune eccellenti pubblicazioni [11], [12], [13], preme qui evidenziare come le principali realizzazioni furono legate allo sviluppo di strumenti e componenti di elevata qualità i quali furono costruiti senza significativi supporti teorici. Inoltre, se si eccettua una importante applicazione di Gates presso il RAE [14], rimasta tuttavia senza ulteriori sviluppi, queste realizzazioni furono ottenute senza l'utilizzazione di modelli del comportamento dinamico dell'aeroplano basati sulle equazioni del moto.

In definitiva, alla fine degli anni trenta, le ricerche sulla dinamica del velivolo proseguivano sulle linee di ricerca tracciate da Bryan, Bairstow e Melvill Jones, gli aeroplani venivano progettati e messi a punto tenendo soprattutto conto dei risultati della sperimentazione in volo e dei primi risultati in termini di qualità di volo, ed alcuni aeroplani volavano in modo automatico utilizzando le versioni più efficienti degli autopiloti nel frattempo sviluppati. Parallelamente si andava sviluppando un'altra attività, quella dei missili autoguidati. In questo settore si stavano ottenendo i primi risultati.

L'esplosione del conflitto mondiale determinò una impressionante

accelerazione di tutte le attività di ricerca e sviluppo in campo aeronautico e nei settori collegati. In questo contesto ebbe luogo una fondamentale svolta nel settore dell'automazione. La forte pressione per lo sviluppo dei servomeccanismi e dei regolatori ad alte prestazioni per svariati usi, quali la frestura servocontrollata dei longheroni delle ali, il controllo dei processi manifatturieri del materiale fissile, la realizzazione di calcolatori per il controllo del tiro e della navigazione, la movimentazione delle antenne radar, portò ad un'estesa utilizzazione delle tecniche di progetto analitiche sviluppate originariamente per gli amplificatori dei sistemi telefonici presso i laboratori della Bell, in modo particolare da Nyquist [15].

Questa imponente attività si svolse coperta da segreto militare e solo dopo la fine del conflitto apparirono in tempi brevi molte pubblicazioni e libri sull'argomento. In questi testi vengono esposte non solo le nuove teorie sulla risposta in frequenza dei sistemi controllati automaticamente, ma anche, connesse a queste, le prestazioni nel dominio del tempo attraverso la corrispondenza tra Funzione di Trasferimento (F.d.T.) e risposta di transitorio come si evidenzia attraverso il metodo della Trasformata di Laplace (T.d.L.). Anche il metodo di Evans [16] sul luogo delle radici, ispirato a considerazioni sul problema del controllo del volo di aeroplani e missili, fu pubblicato poco tempo dopo.

Nello stesso periodo l'aeronautica stava compiendo il grande balzo in avanti: l'enorme estensione degli inviluppi di volo associati all'introduzione di motori a getto, i corrispondenti cambiamenti di configurazione con l'aumento dei carichi alari, la presenza di masse concentrate in fusoliere lunghe e sottili, l'adozione delle ali a freccia nei grandi bombardieri, crearono tuttavia un complesso di problemi nuovi nel comportamento dinamico del velivolo non più affrontabili sulla base degli approcci consolidati per i velivoli della generazione precedente; tipicamente fra questi problemi vi erano quelli legati all'insufficienza di smorzamento nel moto longitudinale di corto periodo e nel dutch roll.

L'utilizzazione delle nuove tecniche della risposta in frequenza e delle F.d.T. per il progetto degli aeroplani e dei loro sistemi di controllo venne significativamente stimolata dall'esigenza di risolvere questi nuovi problemi; la loro applicazione venne discussa in una serie di articoli scientifici e tesi di dottorato a partire dal 1947 e nel 1950 Bollay nel XIV Wright Brothers Lecture [17] riassunse lo stato dell'arte mettendo in evidenza come il metodo della T.d.L., la tecnica della risposta in frequenza, il criterio di stabilità di Nyquist, il metodo del luogo delle radici, i calcolatori analogici, fossero ormai diventati strumenti di lavoro acquisiti per affrontare i problemi del volo controllato.

2. Sviluppi recenti della dinamica del volo controllato

Esempi ed applicazioni

La breve storia innanzi delinata, ha messo in evidenza come le nuove conoscenze scaturite dall'evoluzione della teoria dei controlli abbiano cominciato ad influire profondamente, a partire dalla fine degli anni cinquanta, sugli approcci per lo studio della dinamica del volo controllato.

La loro influenza è destinata a farsi più marcata con il progresso delle conoscenze nella teoria dei controlli. Questo viene fortemente stimolato anche dai sempre più complessi problemi che l'ingegneria aeronautica e la nascente ingegneria spaziale si trovano ad affrontare.

Senza voler trattare in alcun dettaglio l'evoluzione di questi rapporti «virtuosi» tra i due settori, ben documentati in eccellenti rassegne quali ad esempio quella di Bryson del 1984 [18], si riportano nel seguito tipiche applicazioni, per mostrare alcuni passi di questa evoluzione e al tempo stesso evidenziare l'efficacia e l'adeguatezza delle nuove tecniche di analisi nell'affrontare la globalità dei problemi che caratterizzano la dinamica del sistema velivolo-pilota.

2.1. *La risposta ai comandi in ciclo aperto*

L'impiego delle T.d.L. e i concetti di F.d.T. consentono un fondamentale passo in avanti nella interpretazione ed utilizzazione delle equazioni linearizzate del moto e di tutto il patrimonio di risultati sperimentali precedentemente acquisiti. La figura 1 riporta le predette equazioni nelle forme in cui esse si trovano nel citato volume del 1935 di Melvill Jones [5], che riproducono nella sostanza quelle già apparse nel volume di Bairstow del 1920 [4]. In figura 1 compaiono anche le stesse equazioni estratte dal noto testo di McRuer et al. del 1973 [12], che è una evoluzione dei cosiddetti «Northrop Volumes», che agli inizi degli anni cinquanta costituivano una delle fonti più complete sull'uso delle tecniche analitiche nel progetto dei sistemi di controllo automatico del volo.

La sola differenza apparente consiste nell'introduzione, nel testo di McRuer, della T.d.L., ma grazie ad essa è ora possibile interpretare con grande efficacia la risposta dell'aeroplano in ciclo aperto in termini di F.d.T. La struttura di esse, limitatamente alla dinamica longitudinale, è riassunta nella figura 2 sia nella forma completa che nell'approssimazione di corto periodo. Le tre F.d.T. riportate in figura 2 si riferiscono all'incidenza α , all'angolo di inclinazione dell'asse longitudinale del velivolo o angolo di beccheggio θ , ed all'angolo di traiettoria γ e sono relative al comando longitudinale di equilibratore.

Dynamics of the Airplane - B. MELVILL JONES*Symmetric or Longitudinal Equations*

$$\begin{aligned}
 m [\dot{u} + W_1 q + g \cos \Theta_1 \theta] &= X_u u + X_w w + X_q q \\
 m [\dot{w} - U_1 q + g \sin \Theta_1 \theta] &= Z_u u + Z_w w + Z_q q \\
 B \dot{q} &= M_u u + M_w w + M_q q + M_{\dot{w}} \dot{w} + M_0
 \end{aligned}$$

Asymmetric or Lateral Equations

$$\begin{aligned}
 m [\dot{v} + U_1 r - W_1 p - g \cos \Theta_1 \phi] &= Y_v v + Y_p p + Y_r r \\
 A \dot{p} - E \dot{r} &= L_v v + L_p p + L_r r + L_0 \\
 C \dot{r} - E \dot{p} &= N_v v + N_p p + N_r r + N_0
 \end{aligned}$$

Aircraft Dynamics and Automatic Control

D. MCRUER - I. ASHKENAS - D. GRAHAM

Symmetric or Longitudinal Equations

$$\begin{aligned}
 su + W_0 q + g \cos \Theta_0 \theta &= X_u u + X_{\dot{w}} s w + X_w w + X_q q + \Sigma X_{\delta} \delta \\
 sw - U_0 q + g \sin \Theta_0 \theta &= Z_u u + Z_{\dot{w}} s w + Z_w w + Z_q q + \Sigma Z_{\delta} \delta \\
 sq &= M_u u + M_{\dot{w}} s w + M_w w + M_q q + \Sigma M_{\delta} \delta
 \end{aligned}$$

Asymmetric or Lateral Equations

$$\begin{aligned}
 su + U_0 r - W_0 p - g \cos \Theta_0 \phi &= Y_v v + Y_{\dot{v}} s v + Y_p p + Y_r r + \Sigma Y_{\delta} \delta \\
 sp - \frac{I_{xz}}{I_x} sr &= L_v v + L_{\dot{v}} s v + L_p p + L_r r + \Sigma L_{\delta} \delta \\
 sr - \frac{I_{xz}}{I_z} sp &= N_v v + N_{\dot{v}} s v + N_p p + N_r r + \Sigma N_{\delta} \delta
 \end{aligned}$$

Figura 1 - Confronto fra le equazioni del moto linearizzate nella forma usata da Melvill Jones (1935) [5] e da Mc Ruer et al. (1973) [12].

Nella figura 2 sono mostrate le forme tipiche delle suddette F.d.T. con coppie di poli complessi e coniugati nonché la forma che il denominatore D può assumere quando M_u sia negativo e tali che risulti

$$\frac{Z_w M_u}{Z_u M_w} \geq 1$$

(caso del tuck-mode) o quando l'aeroplano sia staticamente instabile. Queste ultime due circostanze si traducono nella presenza nel denominatore di uno zero reale positivo.

$$\frac{\alpha}{\delta_e} = \frac{\frac{Z_\delta}{U_0} \left(s + \frac{1}{T_w} \right) (s^2 + 2 \zeta_w \omega_w s + \omega_w^2)}{(s^2 + 2 \zeta_{sp} \omega_{sp} s + \omega_{sp}^2) (s^2 + 2 \zeta_p \omega_p s + \omega_p^2)} \quad \frac{\alpha}{\delta_e} = \frac{\frac{Z_\delta}{U_0} \left(s + \frac{1}{T_w} \right)}{s^2 + 2 \zeta_{sp} \omega_{sp} s + \omega_{sp}^2}$$

$$\frac{\theta}{\delta_e} = \frac{M_\delta \left(s + \frac{1}{T_{\theta 1}} \right) \left(s + \frac{1}{T_{\theta 2}} \right)}{(s^2 + 2 \zeta_{sp} \omega_{sp} s + \omega_{sp}^2) (s^2 + 2 \zeta_p \omega_p s + \omega_p^2)} \quad \frac{\theta}{\delta_e} = \frac{M_\delta \left(s + \frac{1}{T_{\theta 2}} \right)}{(s^2 + 2 \zeta_{sp} \omega_{sp} s + \omega_{sp}^2) s}$$

$$\frac{\gamma}{\delta_e} = \frac{-\frac{Z_\delta}{U_0} \left(s + \frac{1}{T_{h1}} \right) \left(s + \frac{1}{T_{h2}} \right) \left(s + \frac{1}{T_{h3}} \right)}{(s^2 + 2 \zeta_{sp} \omega_{sp} s + \omega_{sp}^2) (s^2 + 2 \zeta_p \omega_p s + \omega_p^2)} \quad \frac{\gamma}{\delta_e} = \frac{-\frac{Z_\delta}{U_0} \left(s + \frac{1}{T_{h2}} \right) \left(s + \frac{1}{T_{h3}} \right)}{(s^2 + 2 \zeta_{sp} \omega_{sp} s + \omega_{sp}^2) s}$$

$$D(s) = (s^2 + 2 \zeta_{sp} \omega_{sp} s + \omega_{sp}^2) \left(s + \frac{1}{T_{p1}} \right) \left(s + \frac{1}{T_{p2}} \right) \quad \text{Aereo con tuck mode}$$

$$D(s) = (s^2 + 2 \zeta_p \omega_p s + \omega_p^2) \left(s + \frac{1}{T_{sp1}} \right) \left(s + \frac{1}{T_{sp2}} \right) \quad \text{Aereo staticamente instabile}$$

Figura 2 - Struttura tipica di alcune F.d.T. del piano longitudinale, nella forma completa (a sinistra) e nell'approssimazione di corto periodo.

Le successive figure 3 mostrano le risposte ad un comando a gradino di equilibratore per due velivoli B-747 e F-104 e per quest'ultimo in due diversi punti dell'involuppo di volo⁽¹⁾. I diagrammi mostrano sia le risposte nel breve periodo che quelle nel lungo periodo.

⁽¹⁾ Per completezza di rappresentazione, nell'Appendice 1 sono riportate le espressioni dei poli e degli zeri in funzione delle derivate aerodinamiche ed i valori numerici che si riferiscono agli esempi trattati.

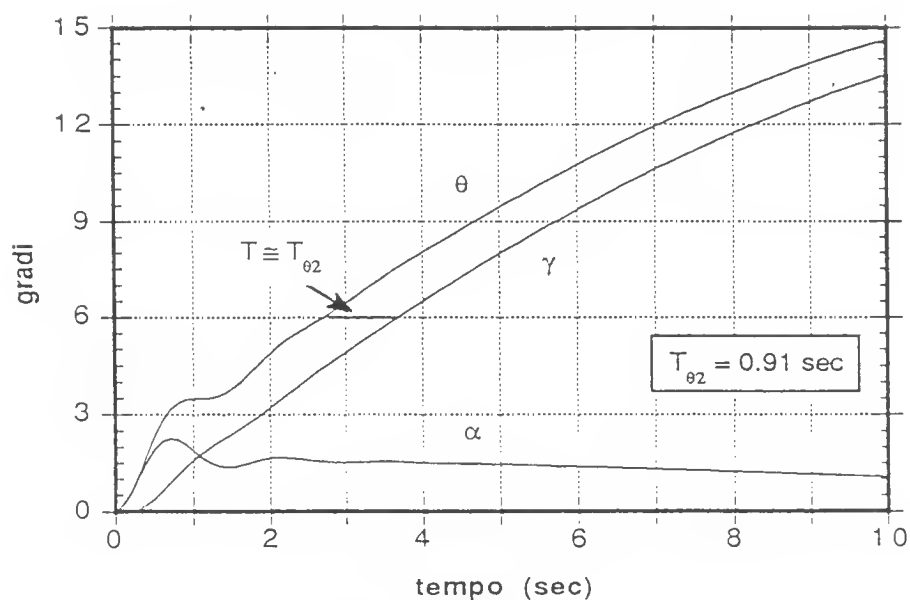


Figura 3a - Risposte in ciclo aperto nel breve periodo, ad un gradino unitario (-1°) di equilibratore. Velivolo F-104 ($M=0.9$, quota = 4500 m).

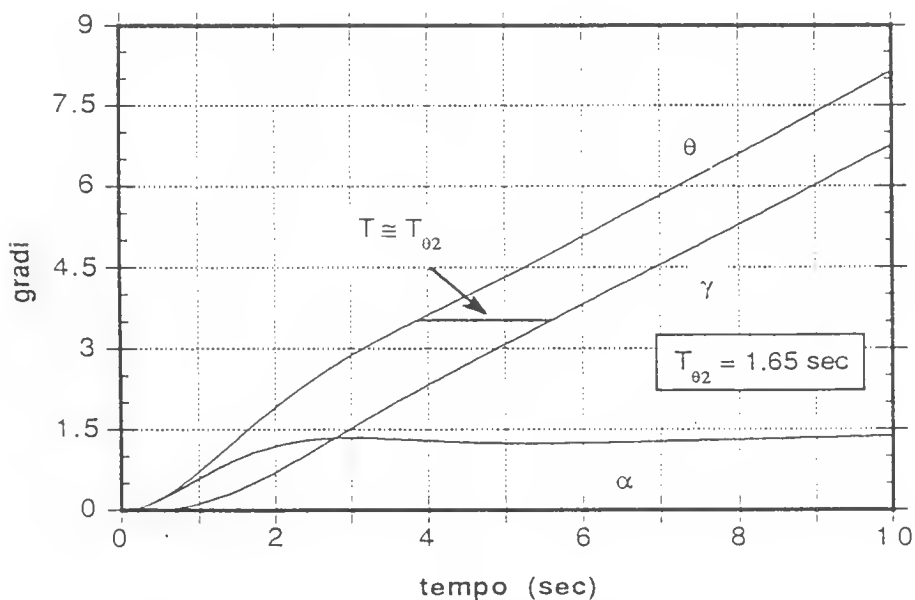


Figura 3b - Risposte in ciclo aperto, nel breve periodo, ad un gradino unitario (-1°) di equilibratore. Velivolo B-747 ($M=0.8$, quota = 6000 m).

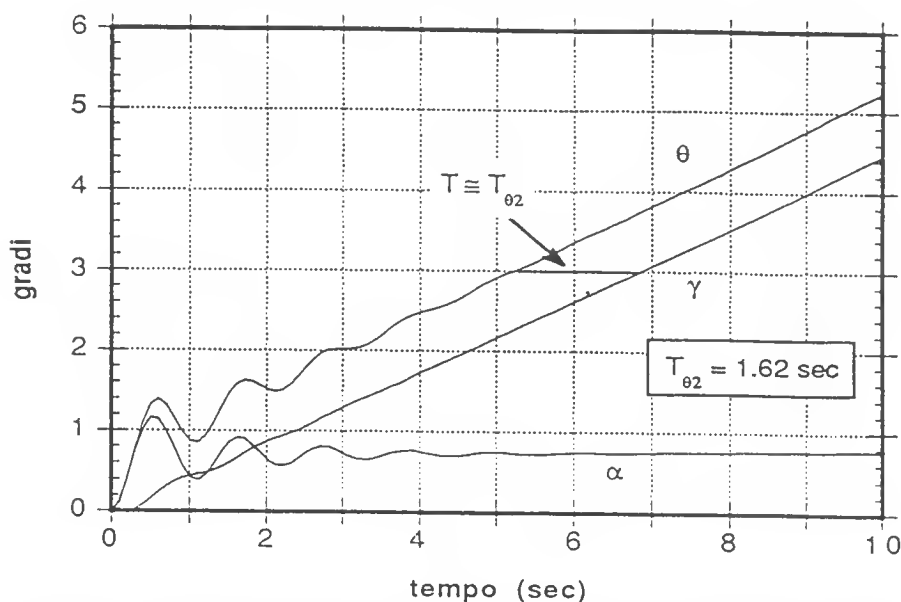


Figura 3c - Risposte in ciclo aperto, nel breve periodo, ad un gradino unitario (-1°) di equilibratore. Velivolo F-104 ($M=1.2$, quota = 10700 m).

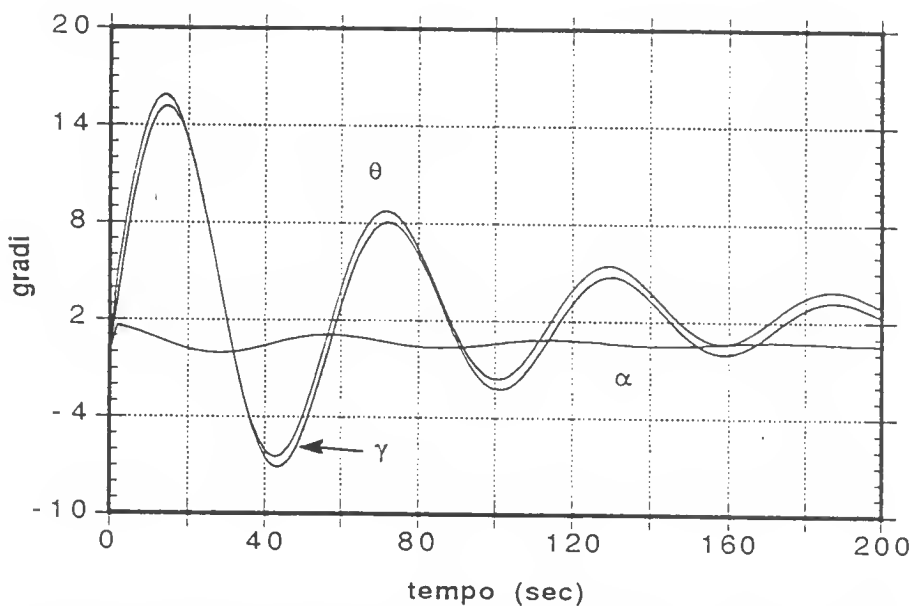


Figura 3d - Risposte in ciclo aperto, nel lungo periodo, ad un gradino unitario (-1°) di equilibratore. Velivolo F-104 ($M=0.9$, quota = 4500 m).

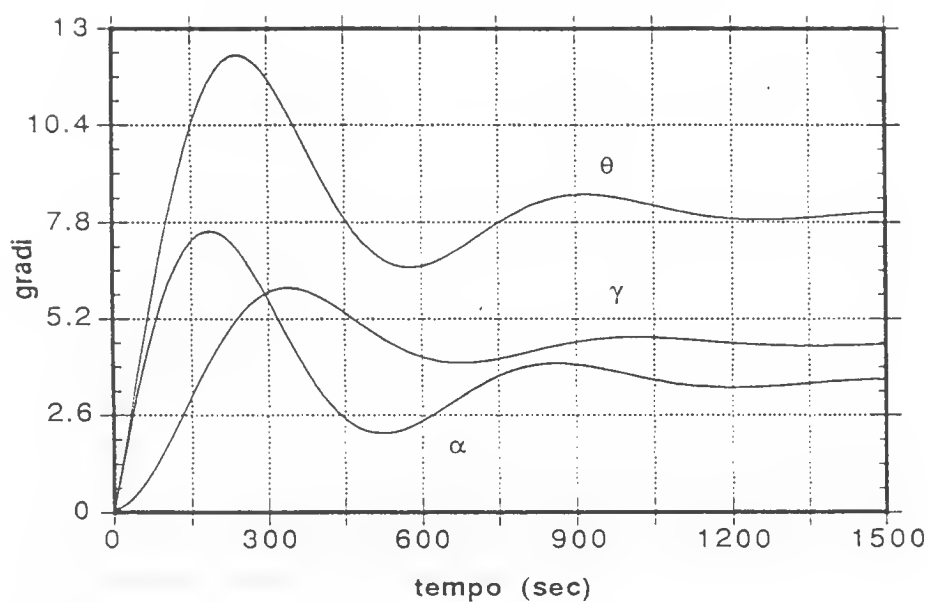


Figura 3e - Risposte in ciclo aperto, nel lungo periodo, ad un gradino di un decimo di grado di equilibratore. Velivolo B-747 ($M=0.8$, quota = 6000 m).

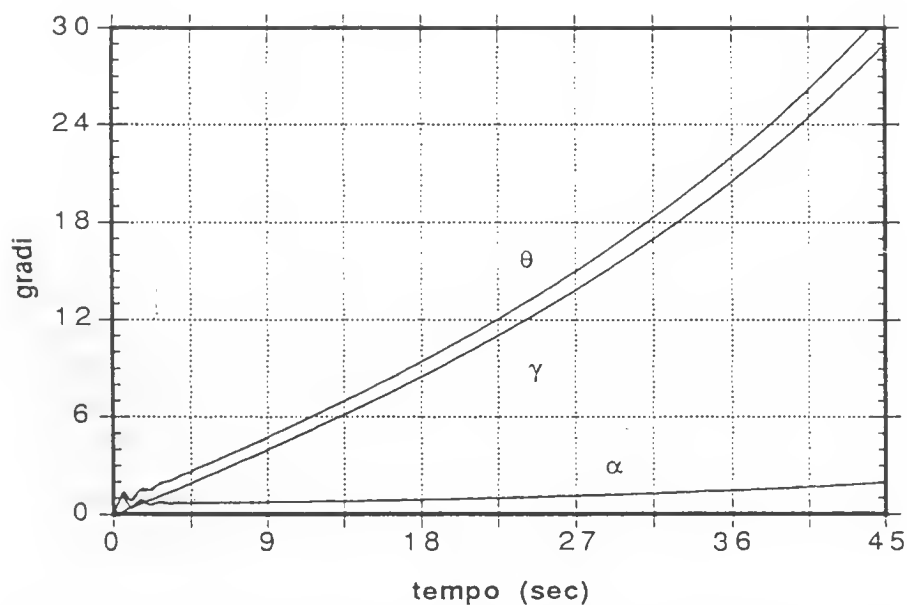


Figura 3f - Risposte in ciclo aperto, nel lungo periodo, ad un gradino unitario (-1°) di equilibratore. Velivolo F-104 ($M=1.2$, quota = 10700 m).

Risposte di questo tipo, pur con notevoli sforzi per l'elevato onere numerico, erano già ottenibili con i metodi classici e di esse si trova traccia nel citato volume di Melvill Jones [5]; tuttavia la loro interpretazione rimaneva oscura e non facilmente comprensibile in termini di proprietà dell'aeroplano e soprattutto di pilotabilità della macchina.

L'interpretazione delle risposte con l'ausilio delle F.d.T. è molto efficace, essendo possibile mettere in relazione le proprietà del moto espresse attraverso i poli, gli zeri ed i guadagni aerodinamici, con le proprietà fisiche dell'aeroplano, descritte dalle derivate aerodinamiche dalle quali poli, zeri e guadagni dipendono funzionalmente secondo le relazioni riportate nell'Appendice 1.

Uno degli aspetti tipici della risposta dell'aeroplano ad un ingresso di equilibratore è il crescere parallelo degli angoli θ e γ . Una volta esauritosi il transitorio in incidenza, l'angolo θ è in anticipo sull'angolo γ di un tempo che, a parte termini trascurabili, si dimostra facilmente essere dato da $T \cong T_{\theta_2}$ e cioè dall'inverso dello zero di alta frequenza della F.d.T. θ/δ_e .

L'angolo θ è facilmente percepibile e misurabile dal pilota, ma è l'angolo γ , non direttamente percepito dal pilota, che definisce la traiettoria. In virtù della differenza costante fra i due angoli, controllando θ si controlla anche γ . Tuttavia è da tener presente che il controllo di γ è influenzato dall'anticipo T_{θ_2} che perciò non deve essere troppo elevato. Ora risulta

$$\frac{1}{T_{\theta_2}} \cong -Z_w = \frac{\rho S U_0 C_{L\alpha}}{2m} = \frac{1}{2} g \frac{U_0 C_{L\alpha}}{W/S}$$

e quindi per i vari tipi di aeroplano in relazione al carico alare W/S , alla velocità di volo U_0 , alla quota di volo (e quindi al valore di densità dell'aria ρ) ed alla forma in pianta dell'ala è possibile trarre rapide conclusioni su questo importante parametro dell'aeroplano. Ad esempio carichi alari elevati, bassi allungamenti alari ed elevate quote di volo tendono ad aumentare il ritardo di γ rispetto a θ peggiorando così le caratteristiche di risposta del velivolo.

Quando si vada ad analizzare la risposta del velivolo tipo F-104 (figura 3d) su tempi molto lunghi — dell'ordine di $(5 \div 10) (2\pi/\omega_p)$, essendo ω_p la pulsazione di lungo periodo — si ritrova la classica dinamica di lungo periodo con le tipiche oscillazioni di θ e γ . Il periodo di oscillazione è dell'ordine di 100 sec, in accordo con la nota conclusione, già insita nell'analisi di Lanchester, che sia

$$\omega_p = \sqrt{-\frac{Z_u g}{U_0}} = \frac{g}{U_0} \sqrt{2}.$$

Il moto è assai poco smorzato come risulta dal numero elevato di oscillazioni presenti prima che sia raggiunta la condizione di regime.

Molto diversa e sostanzialmente atipica è la risposta del velivolo tipo B-747 (figura 3e); il periodo è molto grande, dell'ordine dei 1000 sec., in totale disaccordo con la formula precedente; il moto è anche molto smorzato, altra caratteristica inusuale di questa dinamica.

La spiegazione di questo fatto è da ricercarsi nel ruolo di M_u , la derivata che traduce la variazione di momento di beccheggio al variare di u ; questa variazione è diversa da zero solo se i coefficienti aerodinamici sono influenzati dal numero di Mach, o se sono presenti significativi effetti aeroelastostatici, due fatti entrambi sicuramente presenti in un velivolo di questo tipo. La pulsazione del moto diventa in questo caso

$$\omega_p = \sqrt{-\frac{Z_u g}{U_0} \left[1 - \frac{M_u}{M_w} \frac{Z_w}{Z_u} \right]}.$$

Poiché è $M_u < 0$, cioè l'aeroplano ha tendenza a picchiare in conseguenza di un incremento di velocità u , si ha che

$$0 < \frac{M_u}{M_w} \frac{Z_w}{Z_u} < 1$$

con un valore molto vicini all'unità (nel caso dell'esempio vale 0.962) considerati i parametri dell'aeroplano. Ciò rende il periodo circa dieci volte maggiore del valore predetto dalla formula classica; al tempo stesso anche il fattore di smorzamento risulta, per il basso valore di ω_p , piuttosto elevato, pari a 0.32, in accordo con il numero molto ridotto di oscillazioni rilevabili in figura 3e. Un altro aspetto peculiare rilevabile dalla figura 3e è costituito dalla presenza di forti oscillazioni, nel lungo periodo, nelle variabili α , θ , γ benché l'ampiezza dell'ingresso sia pari solo ad un decimo di grado di rotazione dell'equilibratore. Questa situazione è indicativa di una condizione prossima ad una instabilità di tipo *tuck-mode*. Ad esempio il valore asintotico di α è dato dalla seguente espressione:

$$\alpha = -\frac{M_\delta}{M_\alpha} \frac{1 - \frac{M_u Z_\delta}{M_\delta Z_u}}{1 - \frac{M_u Z_w}{M_w Z_u}}$$

in cui si evidenzia il valore (pari a $-M_\delta/M_\alpha$) che si otterrebbe se fosse trascurabile l'effetto di M_u . Nel caso specifico, il termine legato ad M_u presente al denominatore è, come già in precedenza evidenziato, prossimo all'unità. Di conseguenza il valore di incidenza di regime, che si avrebbe considerando $M_u = 0$, è enormemente amplificato, diventando pari a 34° per un grado di equilibratore. Benché il tempo di crescita del fenomeno descritto sia estremamente lento, permettendo quindi al pilota di controllare l'eccessiva crescita dell'incidenza, si tratta comunque di un esempio di dinamica di lungo periodo non più accettabile in termini di pilotabilità. Un possibile intervento che elimini questo problema, consiste nell'impiego di un sistema di aumento della stabilità basato sulla misura e retroazione della componente longitudinale della velocità del velivolo. Ciò equivale ad utilizzare una legge di controllo che muova l'equilibratore in modo proporzionale al valore di u misurato, ottenendo come risultato quello di modificare il valore di M_u eliminando così la causa del fenomeno descritto.

2.2. Problemi di controllo in ciclo chiuso

Gli esempi citati nel paragrafo precedente e le caratteristiche particolari esaminate, mettono bene in evidenza l'efficacia della rappresentazione della risposta in termini di F.d.T. e spiegano il rapido affermarsi di questo modo di trattare il problema nei primi anni cinquanta.

La descrizione della risposta in termini di F.d.T. trova comunque le sue applicazioni più tipiche e più efficaci in presenza di retroazione. Si è già ricordato che lo sviluppo di nuovi aeroplani fortemente innovativi aveva creato, alla fine degli anni cinquanta, nuovi e complessi problemi nel campo della dinamica del volo. Fra questi aeroplani merita di essere ricordato il B-47; esso per le importanti innovazioni che presentava — ali a freccia, motori in gondole sospesi sotto le ali, quote di volo e numeri di Mach elevati — divenne il capostipite di una lunga generazione di aeroplani progettati per impiego sia civile che militare. Questo aeroplano con la sua configurazione così innovativa, assieme alle molte caratte-

ristiche positive, presentava anche aspetti negativi per la presenza di oscillazioni laterali poco smorzate. Si legge nell'articolo di White [19] «*Early flights indicated a tendency for the aeroplane to "Dutch Roll" at the lower indicated air speeds and at altitude...; the damping in yaw was less than the pilot desired... Elimination of this flight condition has been accomplished by the use of a servomechanism termed a "yaw damper", consisting of a yaw rate gyro pickup that operates the rudder to improve damping in yaw*».

Il problema dell'insufficiente smorzamento nel dutch roll è rimasto poi comune a tutti gli aeroplani successivi che a quella configurazione si rifacevano.

L'esempio trattato si riferisce a un B-747 che del B-47 può considerarsi uno dei discendenti più illustri. La figura 4 mostra uno schema a blocchi del sistema. La misura di velocità angolare r , effettuata dal giroscopio, viene riportata all'attuatore di comando del timone di direzione attraverso il controllore ed il *wash-out* (un filtro passa alto che esclude di fatto la retroazione nelle virate stazionarie). Il pilota è poi libero di intervenire sui comandi di volo. Si tratta di un sistema di retroazione denominato SAS (Stability Augmentation System)⁽²⁾; analoghi sistemi possono essere sviluppati per migliorare lo smorzamento di corto periodo (pitch damper) o lo smorzamento del moto in rollio (roll damper).

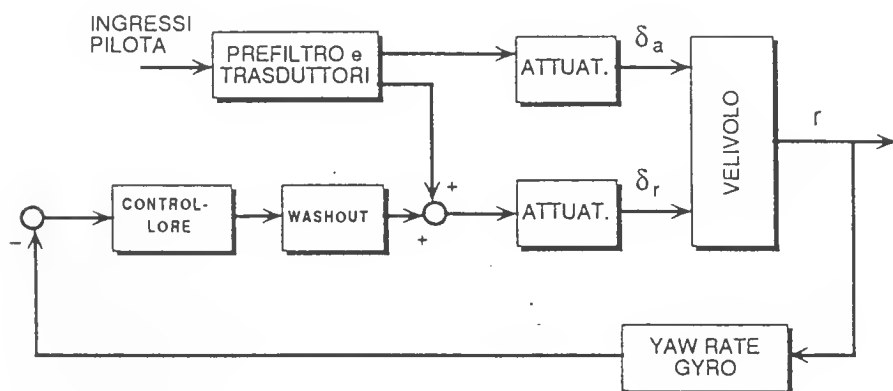


Figura 4 - Schema a blocchi del sistema *yaw damper*.

⁽²⁾ Il nome originale di questi sistemi sviluppati per la prima volta per il B-49-Northrop Flying Wing era «Stability Derivative Augmentation»; il termine Derivative venne poi eliminato per meglio inserire la scritta nell'apposita casella di un disegno di installazione [19], e da allora si parlò di Stability Augmentor o Stability Augmentation System.

Le figure 5.a e 5.b mostrano la risposta in r con e senza *yaw damper* e la relativa escursione del comando. Il problema qui consiste nel determinare una buona combinazione dei valori del guadagno del controllore e della costante del filtro di *wash-out* che porti lo smorzamento del dutch roll a valori accettabili. La scelta dei valori più opportuni di questi parametri dipende dalle caratteristiche dell'aeroplano e dalla sua velocità e quota di volo. Il modo più efficace per risolvere questo tipo di problema è l'uso del luogo delle radici che consente con pochi tentativi di identificare una soluzione soddisfacente.

Il successivo esempio (figura 6), si riferisce allo studio di un autopilota per il controllo dell'assetto longitudinale del velivolo; il loop interno descrive il «*pitch damper*»; il loop esterno introduce un controllo di tipo proporzionale ed integrale. L'obiettivo è quello di selezionare le tre costanti K_θ , K_I e K_q in modo che la variabile in uscita θ segua con sufficiente fedeltà l'ingresso senza eccessive tendenze ad oscillare attorno al valore inseguito.

Utilizzando la tecnica di analisi in frequenza⁽³⁾ le specifiche di sintesi possono essere precisate in termini di banda passante e di margine di fase. Per l'aeroplano utilizzato nell'esempio è possibile scegliere i tre parametri in modo che la banda passante sia dell'ordine della pulsazione del polo di corto periodo e che il margine di fase sia di 60°. Questo garantisce che il sistema in ciclo chiuso si comporti con buona approssimazione come un sistema del 2° ordine rappresentato da una F.d.T. del tipo

$$\frac{\theta}{\theta_i} = \frac{\bar{\omega}_{sp}^2}{s^2 + 2 \bar{\zeta}_{sp} \bar{\omega}_{sp} s + \bar{\omega}_{sp}^2}$$

dove $\bar{\omega}_{sp}$ in pratica coincide con la pulsazione individuata dall'intersezione del diagramma di Bode in ciclo aperto con la linea di 0_{db} e il fattore di smorzamento $\bar{\zeta}_{sp}$, che è controllabile attraverso K_q , è definito in base al valore del margine di fase prescelto. La risposta del sistema è riprodotta nelle figure 7a e 7b nelle quali si trova conferma che il comportamento della risposta in θ è quello anticipato dal modello semplificato del 2° ordine prima delineato. Situazioni di questo tipo si possono

⁽³⁾ Le tecniche di analisi basate sulla risposta in frequenza e sulla costruzione del luogo delle radici, spesso indicate come tecniche «classiche» di controllo, e le loro più significative applicazioni a problemi aeronautici sono prospettate in modo eccellente nel testo di McRuer et al [12].

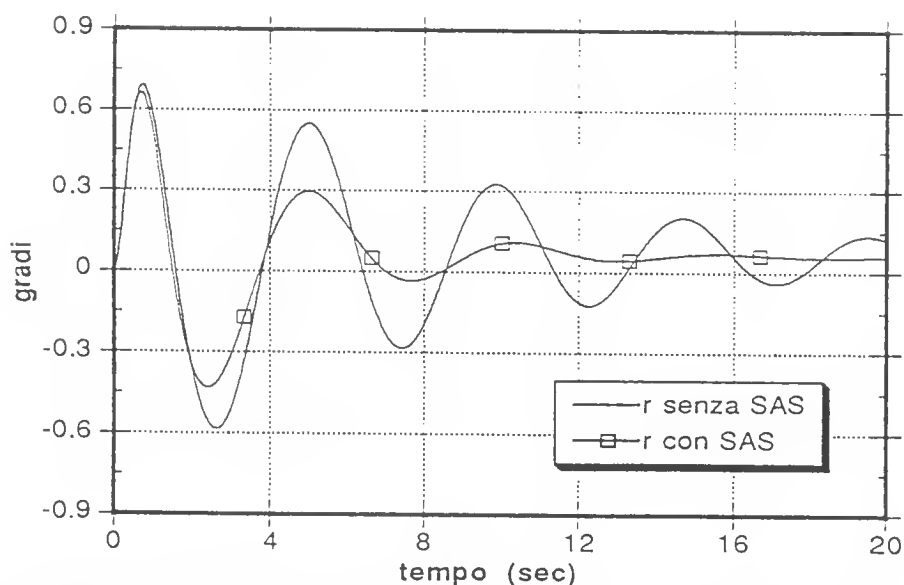


Figura 5a - Risposte in velocità angolare di imbardata r , ad un ingresso impulsivo del timone di direzione. Velivolo B-747 ($M=0.8$, quota = 6000 m).

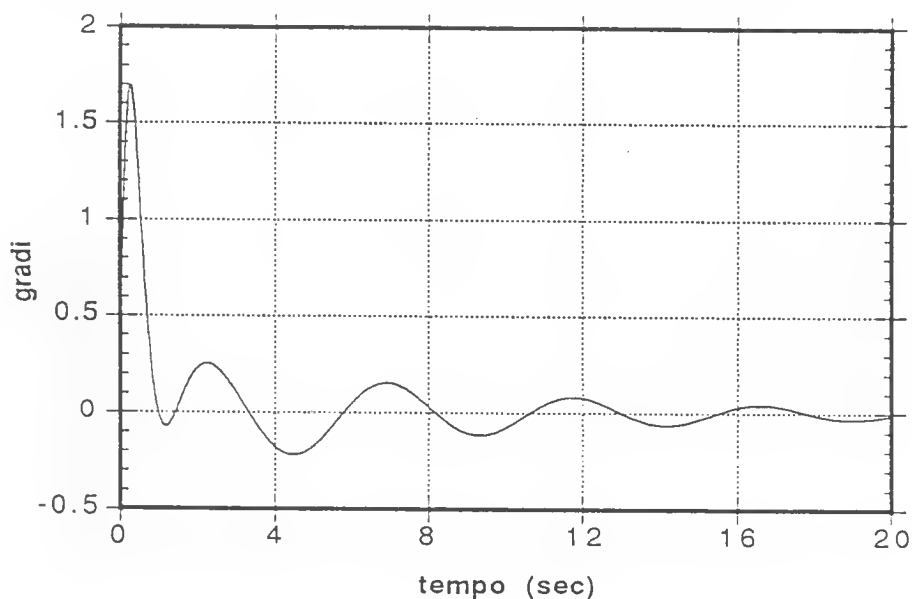


Figura 5b - Escursione del timone di direzione, imposto dallo *yaw damper*, a seguito di un comando impulsivo. Velivolo. Velivolo B-747 ($M=0.8$, quota = 6000 m).

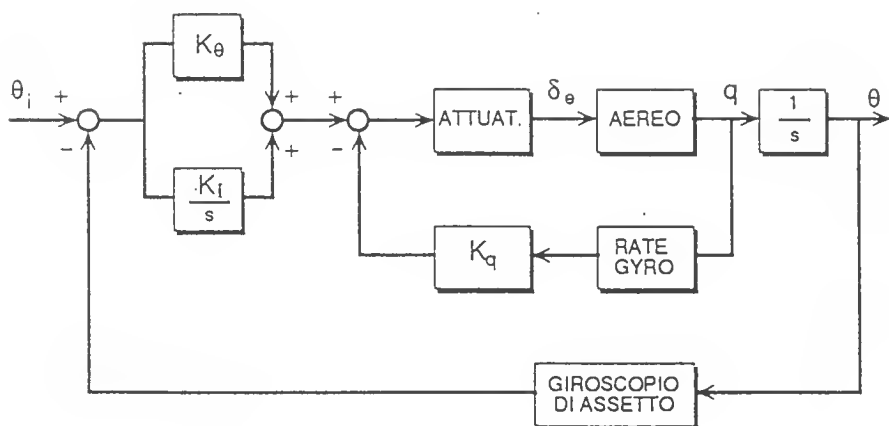


Figura 6 - Schema a blocchi del sistema di autopilota di assetto.

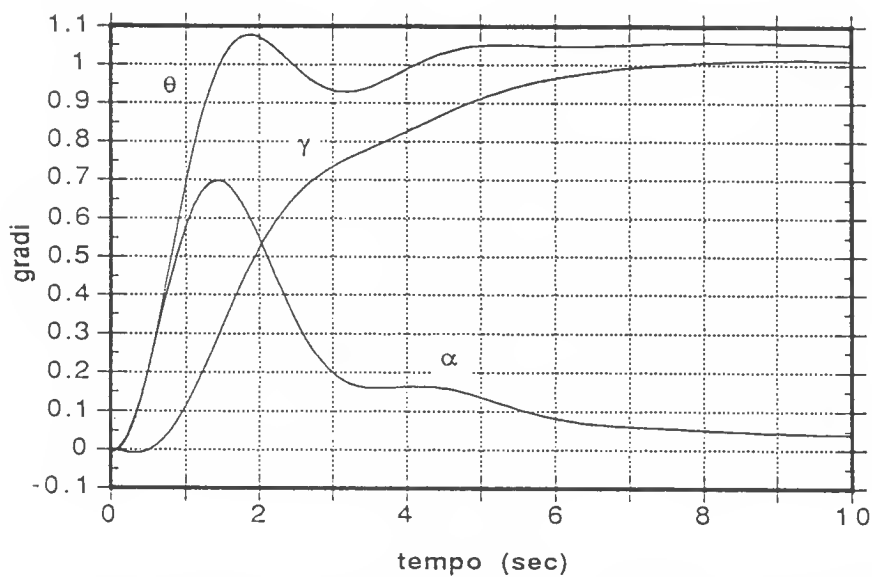


Figura 7a - Risposte ad un ingresso a gradino unitario (1°) di angolo di beccheggio con autopilota di assetto inserito. Velivolo B-747 ($M = 0.8$, quota = 6000m).

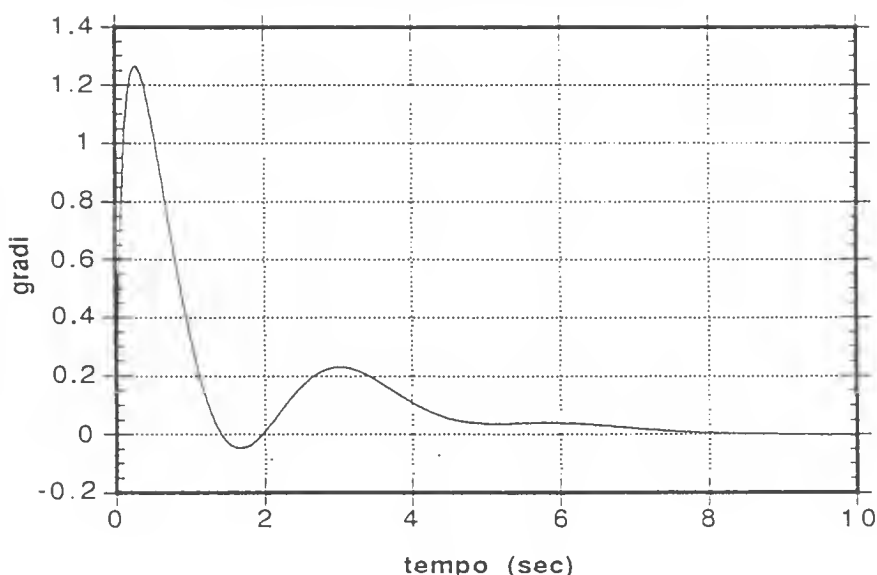


Figura 7b - Escursione dell'equilibratore imposto dal sistema autopilota di assetto con ingresso a gradino unitario. Velivolo B-747 ($M=0.8$, quota = 6000 m).

considerare tipiche per tutti gli aeroplani con una dinamica stabile nel piano longitudinale; anzi nella quasi totalità dei casi si possono ottenere capacità di inseguimento dell'ingresso da parte dell'uscita migliori e con un minor impegno di comando, rispetto a quelle dell'esempio trattato, che sono imputabili alla notevole inerzia dell'aeroplano.

La stessa tecnica di analisi può essere utilizzata per controllare l'assetto di un aeroplano instabile nel piano longitudinale. L'esempio trattato nella figura 8 si riferisce al caso di un aeroplano tipo F-104 nel quale si è supposto di aver modificato il margine di stabilità statica rendendolo negativo. L'uso di un controllo proporzionale e derivativo consente una buona stabilizzazione del sistema come appare dalla figura 9a. Ottenuta la stabilizzazione, il controllo non è stato oggetto di ulteriori affinamenti, che sono probabilmente necessari per migliorare la forma della risposta e l'impegno di comando (figura 9b).

Accanto agli sviluppi fin qui esaminati, basati sulle metodologie dei cosiddetti controlli classici, a partire dagli inizi degli anni '60 vengono sviluppati nuovi metodi, basati sulla descrizione del sistema in termini di variabili di stato e sull'impiego dei metodi variazionali per le definizioni di controllori ottimi. Questa importante parte della teoria dei controlli, denominata, per distinzione con l'altra, teoria moderna, prende

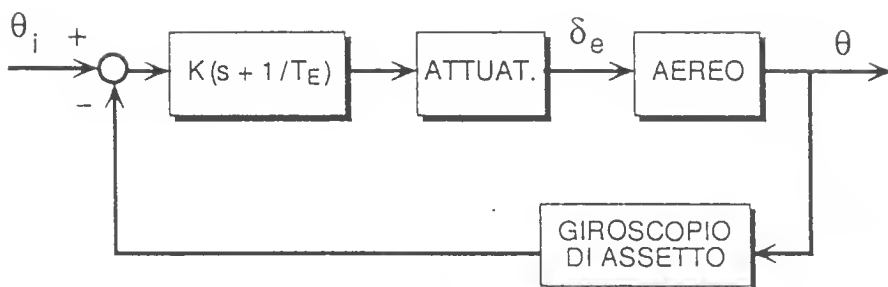


Figura 8 - Schema a blocchi del sistema di autopilota di assetto per un velivolo staticamente instabile.

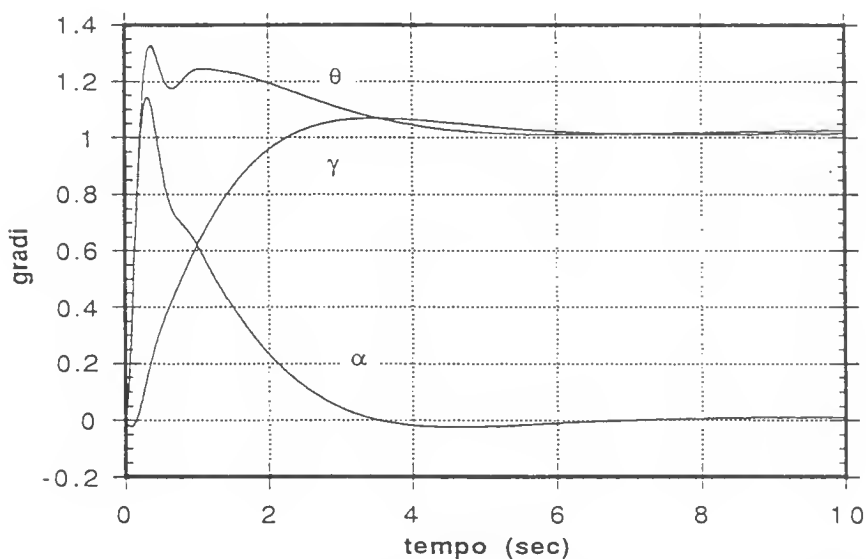


Figura 9a - Risposte ad un ingresso a gradino unitario (1°) di angolo di beccheggio con autopilota di assetto inserito. Velivolo tipo F-104 staticamente instabile.

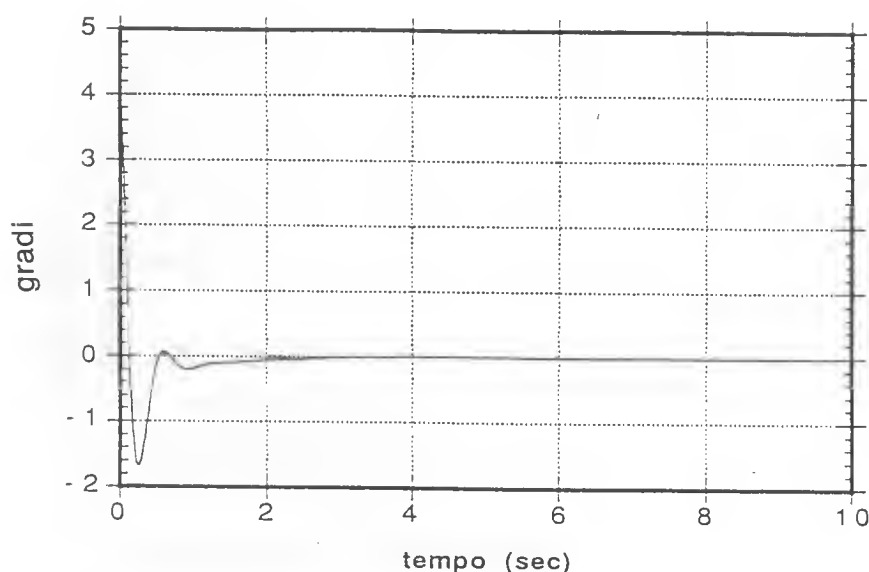


Figura 9b - Escursione dell'equilibratore imposto dal sistema autopilota di assetto con ingresso a gradino unitario. Velivolo tipo F-104 staticamente instabile.

le mosse sotto lo stimolo di specifici problemi applicativi tra cui quelli spaziali, ma rapidamente si sviluppa come branca importante della matematica applicata, rendendo disponibili in tempi brevi una molteplicità di approcci di notevole valore applicativo⁽⁴⁾.

La figura 10 è relativa allo schema a blocchi di un sistema di controllo con retroazione di tutte le variabili di stato, applicato al caso di un velivolo instabile nel piano longitudinale.

L'approccio è quello sviluppato da Kalman [20] per sistemi lineari con indice di merito di tipo quadratico. Più in particolare Kalman ha mostrato che per un sistema dinamico lineare il vettore di controllo «ottimale» è una funzione lineare del vettore di stato; ciò implica che ciascuna delle variabili di stato debba essere portata in retroazione su ciascuna delle variabili di controllo con appropriati guadagni. Con controllo «ottimale» si intende che l'integrale J della forma quadratica nelle variabili di stato e di controllo (figura 11), è minimizzato.

⁽⁴⁾ La letteratura su questi sviluppi è molto ampia ed articolata; il citato articolo di Bryson [18] fornisce un quadro sintetico, ma completo, dei diversi approcci con particolare riferimento ad applicazioni specifiche al campo aeronautico.

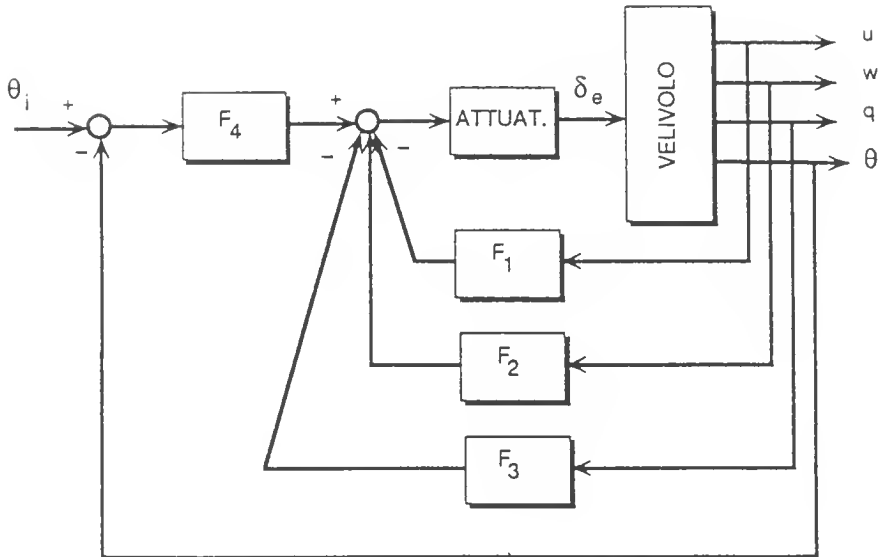


Figura 10 - Schema a blocchi del sistema di controllo dell'assetto longitudinale di un velivolo staticamente instabile, sintetizzato con tecnica LQR.

$$\text{Modello in variabili di stato: } \dot{\underline{x}} = A\underline{x} + B\delta_e \quad \underline{x} = \begin{bmatrix} u \\ w \\ q \\ \theta - \theta_i \end{bmatrix}$$

$$\text{Legge di controllo: } \delta_e = -F_1 u - F_2 w - F_3 q - F_4 (\theta - \theta_i)$$

$$\text{Funzionale quadratico da minimizzare: } J = \int_0^\infty (\underline{x}^T Q \underline{x} + \varrho (\delta_e)^2 dt)$$

$$\text{Valori numerici utilizzati: } Q = \text{diag} (1, 7, 2 \times 10^8, 2 \times 10^8), \quad \varrho = 5 \times 10^8$$

$$\text{Valori calcolati: } F_1 = 0.00013 \text{ rad sec m}^{-1}, \quad F_2 = -0.00066 \text{ rad sec m}^{-1} \\ F_3 = -0.6621 \text{ sec}, \quad F_4 = -0.6319$$

Figura 11 - Formulazione del problema LQR e relativi valori numerici utilizzati.

Questo metodo di sintesi del regolatore del sistema è ora noto come regolatore quadratico lineare (LQR).

Nelle figure 12a e 12b sono mostrate le risposte del sistema e la legge di spostamento del comando.

L'esempio riportato è solo indicativo di una molteplicità di approcci oggi attentamente considerati in campo aeronautico per la sintesi di sistemi di controllo in alternativa e/o in parallelo ai metodi classici i quali, benché le potenzialità dei metodi moderni siano indubbie, continuano a godere di un forte favore nelle aree applicative.

2.3. Esempi di dinamiche del sistema pilota-aeroplano

Nella breve introduzione storica si è ricordato la fondamentale difficoltà incontrata nell'interpretare le proprietà di stabilità e di risposta ai comandi dell'aeroplano in termini di caratteristiche di pilotabilità e qualità di volo.

Si è anche ricordato la via seguita con le sperimentazioni in volo e il progressivo sviluppo di requisiti e specifiche. Uno degli aspetti più importanti delle nuove tecnologie del volo agli inizi degli anni cinquanta è lo sviluppo degli aeroplani a stabilità variabile [6] per le ricerche sulle qualità di volo e la comprensione del ruolo del pilota come parte integrante del sistema.

La figura 13 [6] riproduce un diagramma a blocchi del sistema dinamico pilota-aeroplano. In questo schema il ruolo del pilota appare chiaramente delineato come colui che attua le decisioni sulla base del confronto di ciò che accade e di ciò che si vuole che accada, utilizzando i comandi per introdurre le conseguenti azioni correttive. Quindi il sistema è costituito dall'insieme del pilota e dell'aeroplano i quali agiscono in modo integrato come un sistema in ciclo chiuso la cui dinamica può essere profondamente diversa da quella dell'aeroplano in ciclo aperto che opera senza il continuo intervento del pilota. Questo schema di riferimento consente di interpretare il sistema dando il giusto peso a tutti i componenti, in particolare a quelli che direttamente attengono alle azioni di pilotaggio.

A titolo di esempio, si considerino gli sforzi di barra facendo riferimento ad una manovra di acquisizione di un bersaglio mobile al suolo. Se si immagina lo stesso aeroplano con due diversi gradienti di sforzo di barra, e cioè due diversi «*feel-systems*», il pilota tenderà a considerare l'aeroplano con gradiente minore come più rapido nella risposta, ma con tendenza ad oscillare nell'acquisizione del bersaglio; nell'altro caso descriverà la risposta come più lenta, anche se l'aeroplano tenderà ad acquisire il bersaglio senza oscillazioni.

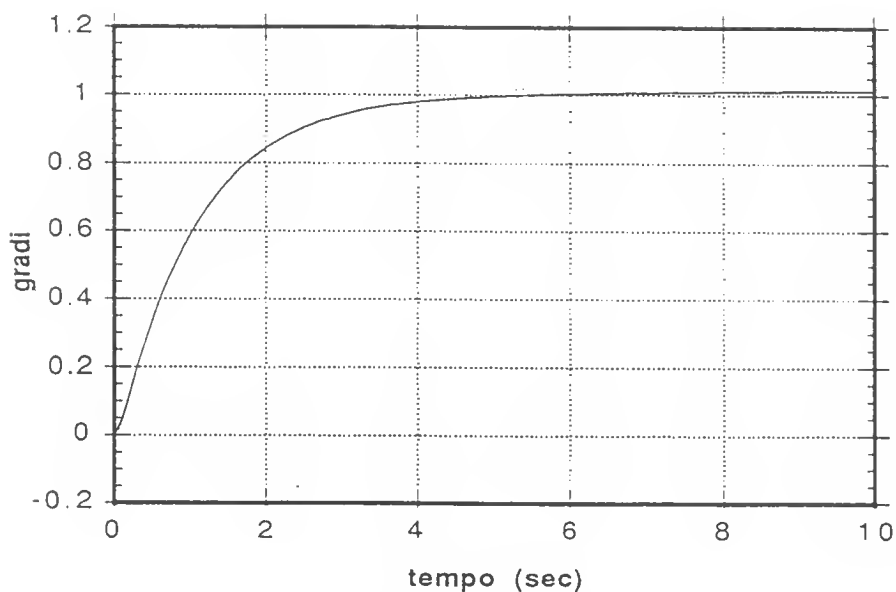


Figura 12a - Risposta in assetto del sistema LQR ad un ingresso in assetto a gradino unitario. Velivolo tipo F-104 staticamente instabile.

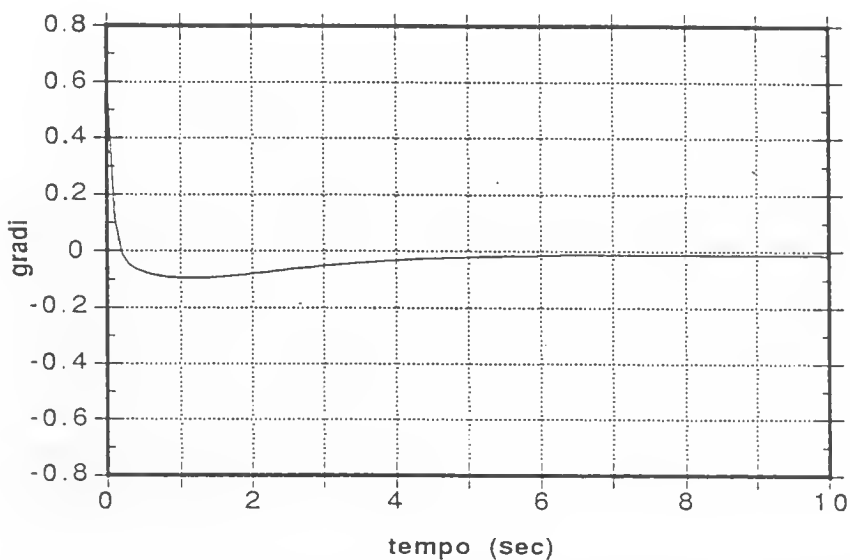


Figura 12b - Escursione dell'equilibratore imposta dal sistema LQR, a seguito di un ingresso in assetto a gradino unitario. Velivolo tipo F-104 staticamente instabile.

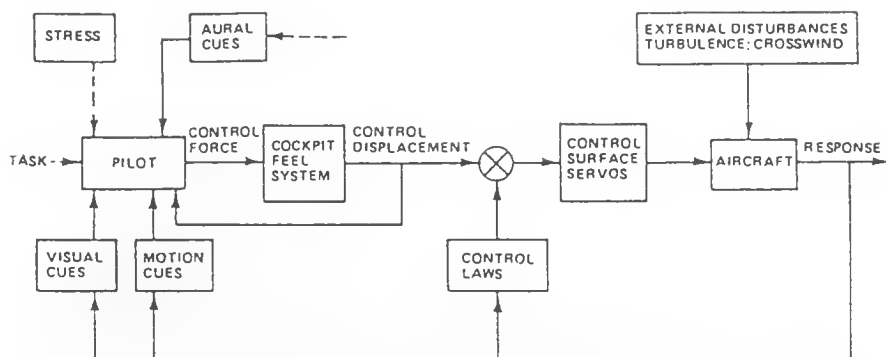


Figura 13 - Schema a blocchi del comportamento del pilota umano come elemento della catena di retroazione [6].

Influenze paragonabili possono avere gli altri elementi del sistema, quali ad esempio i pannelli di presentazione dati (*display*) che con la loro dinamica (movimenti dei simboli, ritardi) condizionano l'interpretazione che il pilota dà dello stato del velivolo e quindi le sue corrispondenti reazioni sui comandi.

L'interpretazione del pilota umano come parte integrante del sistema «macchina volante» ed il successo dei metodi analitici nello studio dell'aeroplano con sistemi di controllo in ciclo chiuso, hanno stimolato la ricerca di modelli matematici capaci di rappresentare almeno gli elementi essenziali del pilota stesso, nei suoi diversi «tasks» di pilotaggio, nella sua funzione di controllore. L'obiettivo è ambizioso ed il compito complesso. Ogni pilota si comporta in base a sue caratteristiche intrinseche, al suo addestramento ed alla sua passata esperienza. Egli inoltre tende a modificare i suoi modi di operare da condizione di volo a condizione di volo e cioè le sue caratteristiche di controllore sono di tipo adattativo sia nei riguardi dell'aeroplano che dei diversi *tasks* di pilotaggio.

Malgrado ciò, il comportamento di un insieme di piloti come elemento del sistema «pilota-aeroplano» tende ad essere sufficientemente omogeneo in presenza di un comune addestramento. Questo fatto, unitamente ai vantaggi derivanti dall'utilizzazione di un approccio analitico al problema, hanno portato allo sviluppo di opportuni modelli matematici che rappresentano il comportamento del pilota mediante un modello linealizzato. La rappresentazione comunemente utilizzata, disponibile già alla fine degli anni sessanta [21], descrive il pilota mediante una F.d.T. di questo tipo

$$Y_p(s) = K_p e^{-\tau s} \frac{T_D s + 1}{T_I s + 1} \left\{ \frac{1}{(T_N s + 1) \left[\left(\frac{s}{\omega_N} \right)^2 + 2 \frac{\zeta_N}{\omega_N} s + 1 \right]} \right\}$$

dove K_p è il guadagno con il quale il pilota chiude il ciclo, di cui esso rappresenta il controllore, secondo lo schema riportato in figura 14 per il controllo longitudinale dell'assetto del velivolo. $e^{-\tau s}$ è il ritardo puro di trasmissione (ritardo fisiologico) legato agli stimoli ed alla conduzione nervosa. La parte tra parentesi graffe descrive la dinamica del sistema neuromuscolare del braccio, mentre la rimanente coppia polo-zero, rappresenta la parte autoadattativa che il pilota è in grado di sviluppare (operando quindi come una rete di anticipo-ritardo) a seconda del particolare velivolo e della condizione di volo in cui si trova ad operare. Valori caratteristici dei parametri che intervengono nel modello matematico sono i seguenti: τ è compreso fra 0.06 e 0.1 sec, $T_N = 0.1$ sec, $\omega_N = 16.5$ rad/sec, $\zeta_N = 0.12$.

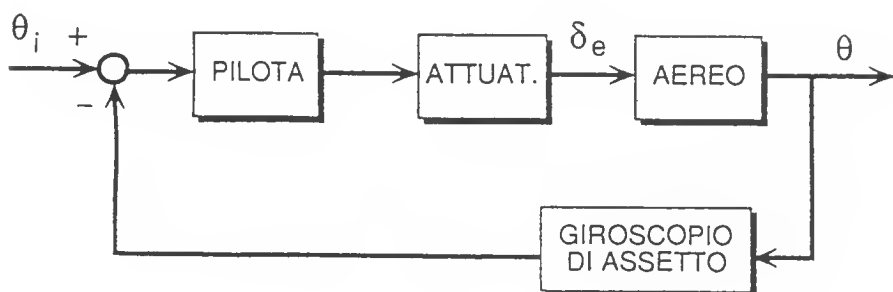


Figura 14 - Schema a blocchi del sistema di autopilota di assetto con il pilota nel *loop* di controllo.

Le regole che devono essere adottate per valutare nella simulazione numerica i valori di T_D e T_I in modo da riprodurre il comportamento del pilota sono state elaborate da McRuer et al. nel citato riferimento [21] e sono sostanzialmente analoghe a quelle impiegate nella sintesi classica dei servomeccanismi, basate sul raggiungimento di una adeguata banda passante del sistema in ciclo chiuso e di un opportuno margine di fase. Un aspetto importante del modello è la sua stretta correlazione con le qualità di volo dell'aeroplano, nel senso che tanto più il pilota tende a

comportarsi come un guadagno puro, tanto migliori sono i suoi giudizi sulle caratteristiche dell'aeroplano per quel particolare task di pilotaggio.

Facendo nuovamente riferimento al problema del controllo dell'assetto longitudinale da parte del pilota, la figura 15 mostra il modo di rispondere del sistema ad un ingresso a gradino di θ_i . Il comportamento è stato analizzato per due diverse leggi di controllo relative a due valori della costante di anticipo T_D al cui aumentare migliora la risposta dell'aeroplano (diventando più rapida e più smorzata), ma aumenta allo stesso tempo la difficoltà della manovra.

Nello stesso diagramma si esamina il ruolo del ritardo fisiologico τ ; il valore 0.1 sec è tipico di un pilota normale in condizioni normali, il valore 0.5 sec può assumersi come indicativo di una condizione fisica notevolmente deteriorata.

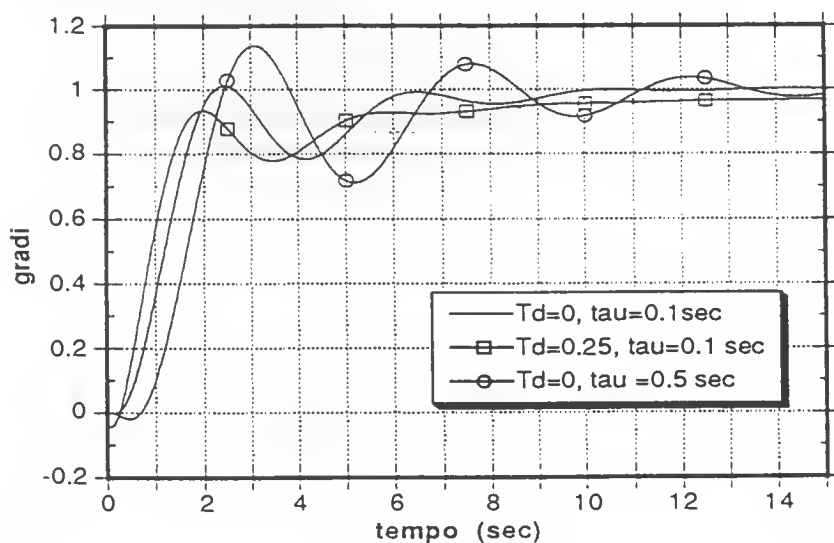


Figura 15 - Risposte in assetto ad un ingresso a gradino di θ_i . T_d è l'anticipo sviluppato dal pilota nella compensazione, τ è il suo ritardo fisiologico. Velivolo B-747.

Si è già messo in evidenza in precedenza come il controllo di θ porti con sé il controllo dell'angolo γ di traiettoria; nella figura 16 compare appunto il modo di variare di γ che conferma quanto anticipato circa il legame fra i due angoli.

Infine nella figura 17 è riportata la legge di spostamento della superficie di controllo e quindi dalla barra di comando nei due casi di assenza

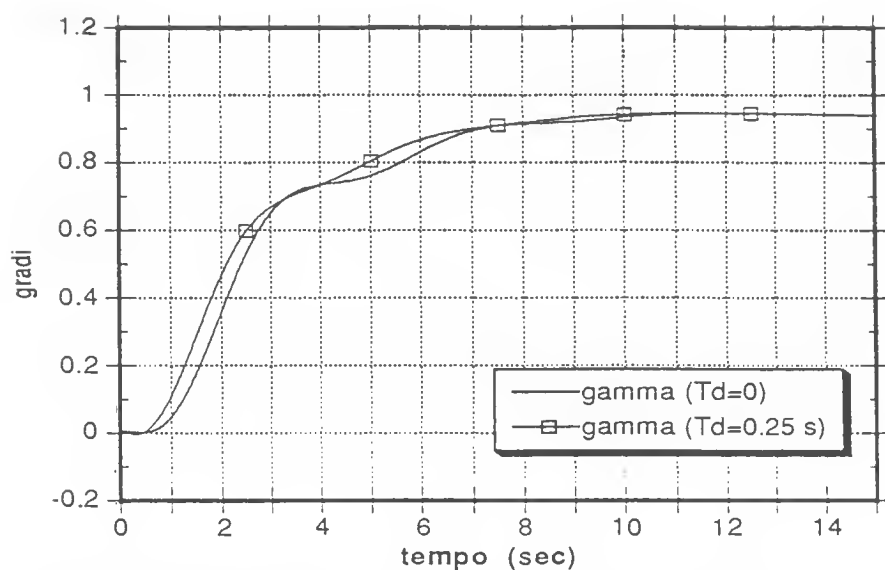


Figura 16 - Risposte in angolo di traiettoria ad un ingresso a gradino di θ_i con il pilota nel *loop*. Nei grafici si assume un ritardo fisiologico di 0.1 sec. Caso B-747.

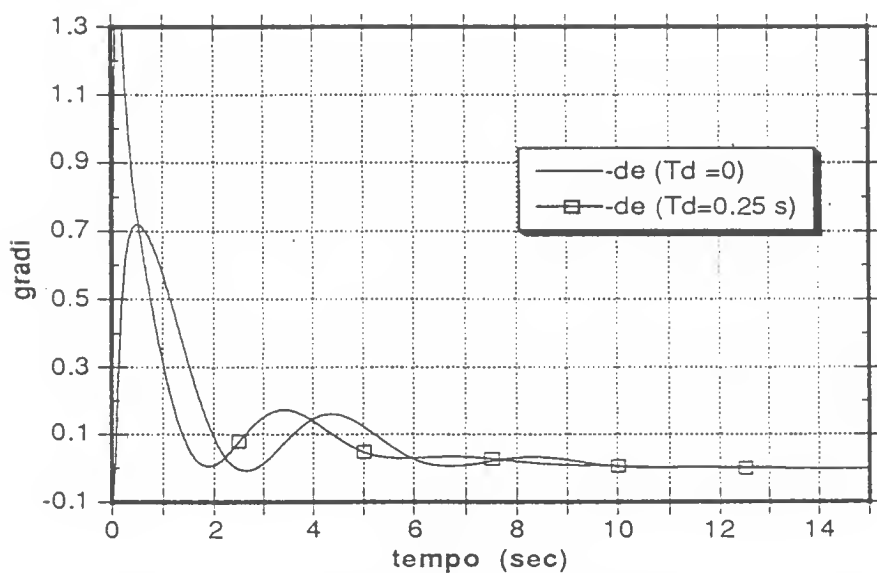


Figura 17 - Escursione dell'equilibratore a seguito dell'inseguimento di un valore di θ_i a gradino con pilota nel *loop*. Ritardo fisiologico = 0.1 sec. Velivolo B-747.

di anticipo o di anticipo $T_D = 0.25$ sec. Appare evidente come l'anticipo generato dal pilota si espliciti con un movimento più rapido e più ampio che non nel caso di semplice controllo proporzionale.

Il secondo schema di controllo analizzato è quello della quota ed è riportato in figura 18. In questo caso il controllo tiene conto di due esigenze concomitanti: l'inseguimento del valore di quota desiderato ed il mantenimento di un adeguato smorzamento nella risposta. Per quest'ultimo obiettivo il pilota tende a chiudere un ciclo su θ con controllo proporzionale e derivativo ($T_D \neq 0$), mentre chiude il loop esterno in modo semplicemente proporzionale. I risultati di questo secondo esempio che descrive un task di pilotaggio piuttosto complesso sono riportati nelle figure 19a e 19b. Esse mettono in evidenza l'efficacia di questo strumento che consente di descrivere situazioni dinamiche molto complesse ed indicare ad un tempo leggi di comando di riferimento per l'effettuazione delle corrispondenti manovre.

Gli esempi trattati sono stati sviluppati per ingressi a gradino della variabile d'ingresso; la generalizzazione ad ingressi caratterizzati da leggi temporali diverse è, dal punto di vista numerico, immediato; l'applicazione ad alcune situazioni particolari può risultare utile, ma in generale l'informazione che si ottiene con l'ingresso a gradino è già di per sé sufficiente per caratterizzare i transitori di risposte con il pilota nel loop.

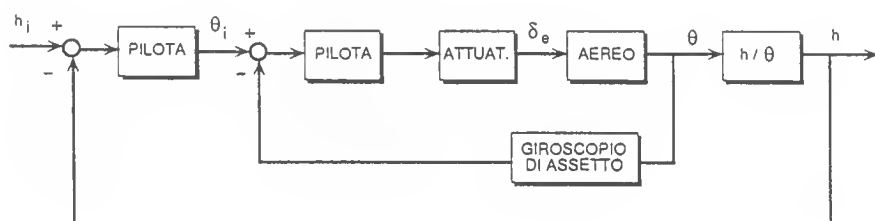


Figura 18 - Schema a blocchi del sistema di autopilota della quota con il pilota nel *loop* di controllo.

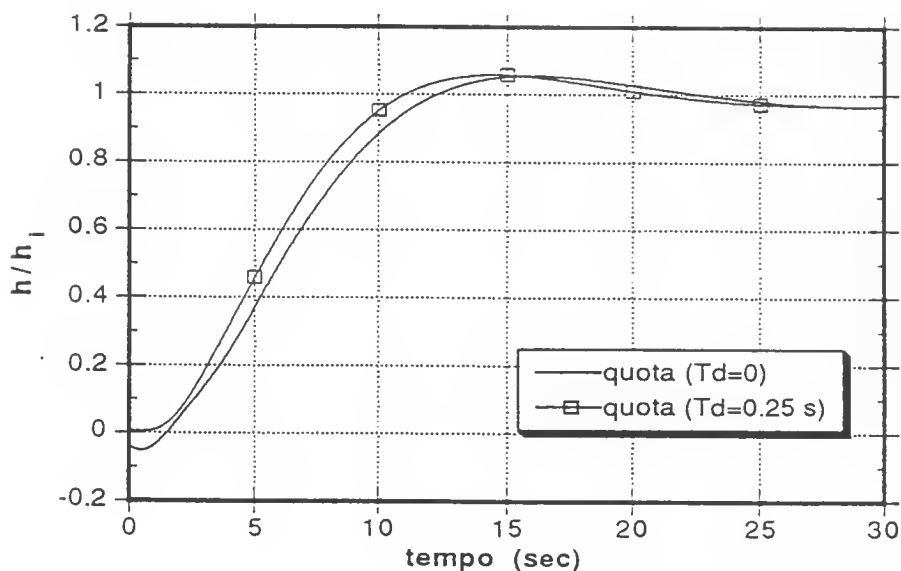


Figura 19a - Risposte in quota ad un ingresso a gradino unitario in quota, con pilota nel loop. Ritardo fisiologico del pilota pari a 0.1 sec. Velivolo B-747.

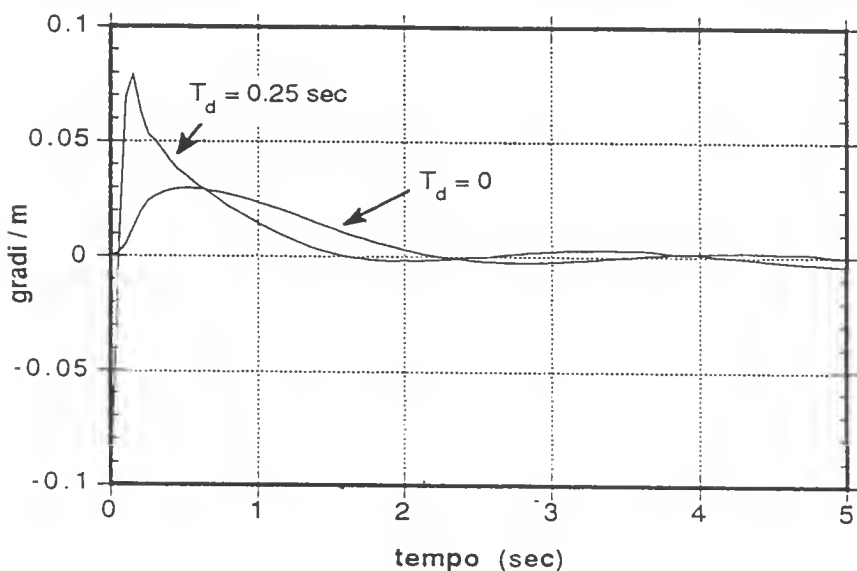


Figura 19b - Escursione dell'equilibratore a seguito dell'inseguimento di un valore di h_i a gradino con pilota nel loop. Ritardo fisiologico = 0.1 sec. Velivolo B-747.

3. Recenti sviluppi e prospettive future

I recenti sviluppi nella tecnologia dei controlli, la disponibilità a bordo di sistemi di calcolo e di elaborazione senza significative limitazioni, ha reso conveniente l'impiego dei sistemi di controllo in aree molto più ampie di quelle tradizionali per ottenere benefici nelle prestazioni e/o riduzioni nei costi.

Le aree tipiche d'intervento riguardano:

- stabilità artificiale,
- alleviazione dei carichi,
- aumento delle velocità di flutter,
- controllo della posizione del baricentro,
- limitazioni all'inviluppo di volo,
- miglioramento del livello di comfort di passeggeri ed equipaggio,
- riduzione dell'affaticamento delle strutture.

I vantaggi conseguibili in termini di prestazioni e/o costi vengono comunque pagati in termini di complessità dei sistemi a causa del numero elevato dei sottosistemi che è necessario prevedere.

Un tipico esempio di questo stato di cose si trova nella figura 20 che confronta la F.d.T. del velivolo convenzionale nella risposta di corto periodo con quello di un moderno velivolo da combattimento con sistemi di controllo attivo complessi.

Uno dei problemi che questa nuova situazione pone è l'interpretazione di queste macchine in termini di qualità di volo e quindi di pilotabilità dell'aeroplano. È chiaro infatti che i punti fermi raggiunti in termini di qualità di volo, codificati dalle MIL-F-8785-C perdono di significato di fronte a questi sistemi.

Si è aperta quindi una nuova era nello studio e nella interpretazione della dinamica del volo controllato. Molti nuovi problemi sono sul tappeto e di questo stato di cose fa fede il nuovo modo di concepire il problema delle qualità di volo, con l'abbandono della MIL predetta e l'adozione della attuale versione delle specifiche di qualità di volo definite nella MIL-STD-1797A pubblicata nel 1990 e del relativo manuale interpretativo.

Più in generale, l'imponente espansione in termini di hardware-software che ha avuto luogo nell'ultimo decennio e la conseguente progressiva acquisizione della capacità di usare i controlli di volo per configurare il velivolo fin dalle prime fasi di progetto, con interventi specifici nelle aree prima delineate, comportano uno straordinario incremento nella complessità del sistema ed una nuova difficoltà culturale di dominarne i comportamenti. Questo anche in presenza dello straordinario aumento

della capacità di calcolo che si è venuta sviluppando, cui spesso non fa riscontro una effettiva capacità di dominare e quindi modellare adeguatamente il sistema in tutte le sue parti.

Un nuovo sforzo di integrazione di diverse culture dell'ingegneria, quali le tecnologie del volo, i controlli, la componentistica avanzata, i sistemi hardware-software di bordo, è oggi necessario per riguadagnare il pieno dominio del sistema aeroplano. Questo sforzo d'integrazione è la nuova sfida culturale che sta alla base dei nuovi sviluppi dell'aeronautica.

La sfida è in atto...

$$\frac{\dot{\theta}}{F_s} = \frac{K (s + 1/T_{\theta_2})}{s^2 + 2\zeta\omega s + \omega^2} = \frac{(1/T_{\theta_2})}{[\zeta; \omega]}$$

$$\frac{\dot{\theta}}{F_s} = \frac{[-0.5; 127] [-0.3; 83] [-0.07; 61] [1; 0.9] [1; 1] [1; 1.2] [1; 2.3] [1; 2.8] (4) [0.1; 53] (13) (14)}{(0.5) (0.8) (0.9) [1; 1.1] (2) [1; 2.4] [0.7; 4.1] (3.4) (3.7) (6.7) [0.2; 53] [1; 13] [0.4; 33] [0.7; 20]} \dots$$

$$X \frac{[0.4; 34] [0.34; 46] [0.55; 39] [1; 22] [0.7; 37] [0.9; 34] [0.9; 34] (35)}{[0.3; 42] [0.9; 16] [0.6; 27] [0.3; 66] [0.4; 85] [0.4; 105] [0.4; 105] [0.9; 47] [0.3; 140] [0.4; 180]} \dots$$

$$X \frac{1}{[0.7; 102] (72) [0.5; 228] (112) [0.6; 279] [0.8; 329] (259) [0.9; 278] [1; 410]} \dots$$

X Terms due to digital system

Figura 20 - Confronto fra le F.d.T. in velocità angolare di beccheggio fra un velivolo convenzionale ed un moderno velivolo da combattimento.

BIBLIOGRAFIA

- [1] HOWARD R.W., «*Automatic Flight Controls in Fixed Wing Aircraft. The First 100 Years*» - Aeronautical Journal, Novembre 1973, pp. 533-562.
- [2] BRYAN G.H., WILLIAMS W.E., «*The Longitudinale Stability of Aerial Gliders*» - Proc. Roy. Soc. Series A, 73, no. 489, 1904, pp. 100-116.
- [3] BAIRSTOW L., MELVILL JONES B., THOMPSON B.A., «*Investigation into the Stability of an Airplane*» - ARC R & M 77, 1913.
- [4] BAIRSTOW L., «*Applied Aerodynamics*» - Longmans Green, London, 1920.
- [5] MELVILL JONES B., «*Dynamics of the Airplane*», In «*Aerodynamic Theory*» by Durand W.F. Editor in Chief, Volume V, Julius Springer, Berlin, 1935, pp. 1-222.
- [6] HARPER Jr. R.P., COOPER G.E., «*Handling Qualities and Pilot Evaluation*», Journal of Guidance, Control, and Dynamics, Vol. 9 no. 5 Spet-Oct. 1986, pp. 515-529.
- [7] PERKINS C.D., «*Development of Airplane Stability and Control Technology*», Journal of Aircraft, Vol. 7 no. 4 July-August 1970, pp. 290-301.
- [8] GILRUTH R.R., «*Requirements for Satisfactory Flying Qualities of Airplanes*», NACA TR 775, 1943.
- [9] Anon. «*Stability and Control Requirements for Airplanes*», AAF-C-1815, 1943.
- [10] LEGGETT D.B., BLACK G.T., «*MIL-STD-1797 Is Not a Cookbook*», Flying Qualities, AGARD-CP-508, Feb. 1991, pp. 7.1-7.19.
- [11] MCRUER D., GRAHAM D., «*A Historical Perspective for Advances in Flight Control Systems*», Advances in Control Systems, AGARD-CP-137, May 1974, pp. 2.1-2.7.
- [12] MCRUER D., ASHKENAS I., GRAHAM D., «*Aircraft Dynamics and Automatic Control*», Princeton University Press, 1973.
- [13] MCRUER D., GRAHAM D., «*Eighty Years of Flight Control: Triumphs and Pitfalls of the Systems Approach*», Journal of Guidance, Control, and Dynamics, Vol. 4 no. 4 July-August 1981, p. 353-362.
- [14] GATES S.B., «*Notes on the Aerodynamics of Automatic Directional Control*», RAE Rept. no. BA 487, Feb. 19, 1924; «*Notes on the Aerodynamics of an Altitude Elevator Control*», RAE no. BA 494, Mar. 19, 1924.

- [15] NYQUIST H., «*Regeneration Theory*», Bell Systems Tech. J., 11, no. 1, Jan. 1932, pp. 126-147.
- [16] EVANS W.R., «*Graphical Analysis of Control Systems*», Trans. AIEE, 67, 1948, pp. 547-551.
- [17] BOLLAY W., «*Aerodynamic Stability and Automatic Control*», Journal of the Aeronautical Sciences, Vol. 18, no. 9, September 1951, pp. 569-624.
- [18] BRYSON A.E., «*New Concepts in Control Theory, 1959-1984*», Journal of Guidance, Control, and Dynamics, Vol. 8 no. 4, July-August 1985, pp. 417-425.
- [19] WHITE R.J., «*Investigations of Lateral Dynamic Stability in the XB-47 Airplane*», Journal of the Aeronautical Sciences, Vol. 17, no. 3, March 1950, pp. 133-148.
- [20] KALMAN R.E., «*When is a Linear Control System Optimal?*», Journal of the Aeronautical Sciences, Vol. 17, no. 3, March 1950, pp. 133-148.
- [21] MCRUER D., GRAHAM D., KRENDEL E., REISNER JR W., «*Human Pilot Dynamics in Compensatory Systems*», USAF, AFFDL-TR-65-15, 1965.

APPENDICE 1. Espressioni di alcuni poli e zeri delle F.d.T. della dinamica longitudinale e valori numerici utilizzati negli esempi.

Si riportano di seguito le espressioni dei poli e degli zeri delle F.d.T. in α , θ e γ relative ad un ingresso di equilibratore δ_e in funzione delle derivate aerodinamiche di stabilità. La struttura di dette F.d.T. è riassunta nella figura 2.

$$\omega_{sp}^2 \cong -M_\alpha + Z_w M_q$$

$$2 \zeta_{sp} \omega_{sp} \cong -(Z_w + M_q + M_\alpha)$$

$$\omega_p^2 \cong -\frac{Z_u g}{U_0} \left[1 - \frac{M_u}{M_w} \frac{Z_w}{Z_u} \right]$$

$$2 \zeta_p \omega_p \cong -X_u + \frac{M_u (X_\alpha - g)}{M_\alpha}$$

$$\frac{1}{T_{\theta_1}} \cong -X_u + \frac{X_w}{Z_w} Z_u \frac{1 - (Z_\delta/M_\delta) (M_u/Z_u)}{1 - (Z_\delta/M_\delta) (M_w/Z_w)}$$

$$\frac{1}{T_{\theta_2}} \cong -Z_w + \frac{M_w}{M_\delta} Z_\delta$$

$$\frac{1}{T_w} \cong \frac{U_0 M_s}{Z_\delta} \quad \omega_w^2 \cong \frac{g}{U_0} \left(-Z_u + \frac{Z_\delta}{M_\delta} M_u \right)$$

$$2 \zeta_w \omega_w \cong -X_u$$

$$\frac{1}{T_{h_1}} \cong -X_u + (X_\alpha - g) \frac{Z_u}{Z_\alpha} \frac{1 - (Z_\delta/M_\delta) (M_u/Z_u)}{1 - (Z_\delta/M_\delta) (M_w/Z_w)}$$

$$\frac{1}{T_{h_2}} \cong -\frac{1}{T_{h_3}} \cong \sqrt{M_\alpha - \frac{M_\delta}{Z_\delta} Z_\alpha}$$

Per completezza, si riportano le espressioni delle derivate aerodinamiche di stabilità in funzione della densità ρ , della superficie alare S , della velocità di volo U_0 , della massa m , della corda media aerodinamica c , del momento di inerzia attorno all'asse y baricentrico I_y e dei coefficienti di resistenza e di portanza C_D e C_L .

$$X_u = -\frac{\rho S U_0}{m} (C_D + C_{Du})$$

$$X_w = \frac{\rho S U_0}{2m} (C_L - C_{D\alpha})$$

$$Z_u = -\frac{\rho S U_0}{m} (C_L + C_{Lu})$$

$$Z_w = -\frac{\rho S U_0}{2m} (C_{L\alpha} + C_D)$$

$$M_u = \frac{\rho S U_0 c}{I_y} (C_m + C_{mu})$$

$$M_w = \frac{\rho S U_0 c}{2 I_y} C_{m\alpha}$$

$$M_w = \frac{\varrho S c^2}{4 I_y} C_{m\dot{\alpha}}$$

$$M_q = \frac{\varrho S U_0 c^2}{4 I_y} C_{mq}$$

$$Z_d = - \frac{\varrho S U_0^2}{m} C_{L\delta}$$

$$M_\delta = \frac{\varrho S U_0^2 c}{2 I_y} C_{m\delta}.$$

I valori numerici utilizzati negli esempi svolti sono riassunti nel seguente prospetto

Quantità	Unità	F-104 M=0.9, h=4500 m	F-104 M=1.2, h=10700 m	B-747 M=0.8, h=6000 m
U_0	m sec ⁻¹	290	355	253
X_u	sec ⁻¹	-0.0193	-0.0139	-0.0064
X_w	sec ⁻¹	-0.0110	-0.0308	0.0263
Z_u	sec ⁻¹	-0.1207	-0.0269	-0.0941
Z_w	sec ⁻¹	-1.2170	-0.7931	-0.624
Z_δ	rad ⁻¹ sec ⁻¹	-209.7	-144.2	-32.7
M_u	m ⁻¹ sec ⁻¹	0.0144	-0.00810	-0.00072
M_w	m ⁻¹ sec ⁻¹	-0.0669	-0.0925	-0.00501
$M_{\dot{w}}$	m ⁻¹	-0.00156	-0.00042	-0.000695
M_q	sec ⁻¹	-0.956	-0.493	-0.668
M_δ	rad ⁻¹ sec ⁻²	-33.6	-23.3	-2.08

Si riportano inoltre i valori numerici relativi ai poli ed agli zeri

Quantità	Unità	F-104 M=0.9, h=4500 m	F-104 M=1.2, h=10700 m	B-747 M=0.8, h=6000 m
ω_{sp}	rad sec ⁻¹	4.53	5.77	1.3
ζ_{sp}		0.288	0.125	0.576
ω_p	rad sec ⁻¹	0.11	(0.0388)*	0.0098
ζ_p		0.121	(-0.0299)	0.323
$1/T_{\theta_1}$	sec ⁻¹	0.0177	0.0133	0.0105
$1/T_{\theta_2}$	sec ⁻¹	1.094	0.619	0.606
$1/T_W$	sec ⁻¹	153.5	189.0	52.7
ω_w	rad sec ⁻¹	0.07	0.0179	0.0593
ζ_w		0.0096	0.0067	0.0537
$1/T_{h_1}$	sec ⁻¹	0.0133	0.0127	0.0045
$1/T_{h_2}$	sec ⁻¹	11.69	11.01	6.11
$1/T_{h_3}$	sec ⁻¹	-14.61	-10.50	-5.16

* I numeri fra parentesi si riferiscono al fatto che i poli sono reali e non complessi coniugati.



Aircraft Survival in Microburst Encounters

Angelo MIELE*

Abstract. *Windshears originating from downbursts have caused many aircraft accidents in the last two decades. To offset the windshear threat, both avoidance systems and recovery systems have been intensely studied.*

This paper summarizes recent efforts on windshear recovery systems, designed to increase the survival capability of an aircraft. First, optimal trajectories are discussed. Then, guidance schemes approximating in real time the key properties of the optimal trajectories are presented. Future research directions are discussed.

Key Words. *Downbursts, microbursts, windshears, aircraft safety, survival capability, guidance, control, optimal control, optimal trajectories, sequential gradient-restoration algorithm.*

Notations

- D = drag, lb;
- h = altitude, ft;
- L = lift, lb;
- m = mass, lb ft⁻¹ sec²;
- t = running time, sec;
- T = thrust, lb;
- V = relative velocity, fps;
- W = weight, lb;
- W_h = vertical component of wind velocity, fps;
- W_x = horizontal component to wind velocity, fps;
- x = horizontal distance, ft;
- α = relative angle of attack, rad;
- β = engine power setting;

This work was supported by the Aviation Research and Education Foundation and by Texas Advanced Technology Program Grant No. TATP-003604020.

The author is indebted to Mr. G.D. Wu for considerable technical assistance.

* A.J. Foyt Family Professor in Engineering and Professor of Aerospace Sciences/Mathematical Sciences, Aero-Astronautics Group, Rice University, Houston, Texas.

- γ = relative path inclination, rad;
- γ_e = absolute path inclination, rad;
- δ = thrust inclination, rad;
- δ_F = flap deflection, rad;
- λ = wind intensity parameter;
- θ = pitch attitude angle, rad;
- τ = final time, sec.

Superscripts

- $\cdot$ = derivative with respect to time;
- $\sim$ = nominal condition

Subscripts

- 0 = initial point;
- τ = final point.

1. Introduction

The study of strategies for aircraft flight in windshear has been an important research topic for the past decade. Indeed, a windshear encounter can be a serious threat to the safety of flight. According to records covering a 20-year period, some 30 airplane accidents in take-off or landing can be attributed to windshear [Ref. 1]. It is of interest to note that the majority of these accidents took place in either regions of high-density traffic or tropical and subtropical regions. See Fig. 1, which is adapted from Ref. 2.

The cause of the above accidents is a particular meteorological condition, called the downburst by Fujita [Refs. 3-4]. In a downburst, there is a descending column of air, which then spreads horizontally in the neighborhood of the ground. Hence, there is shear coupled with downdraft and updraft (Figs. 2-4).

Fujita classified downbursts in two categories: macrobursts (large-scale downbursts, up to 250 miles) and microbursts (small-scale downbursts, up to 3 miles). Microbursts are more dangerous than macrobursts, because the wind gradients are usually larger.

When an aircraft in take-off or landing enters a microburst, two phenomena affecting flight performance are present: the transition from headwind to tailwind and the downdraft; usually, the former is predominant with respect to the latter. Associated with the headwind-to-tailwind

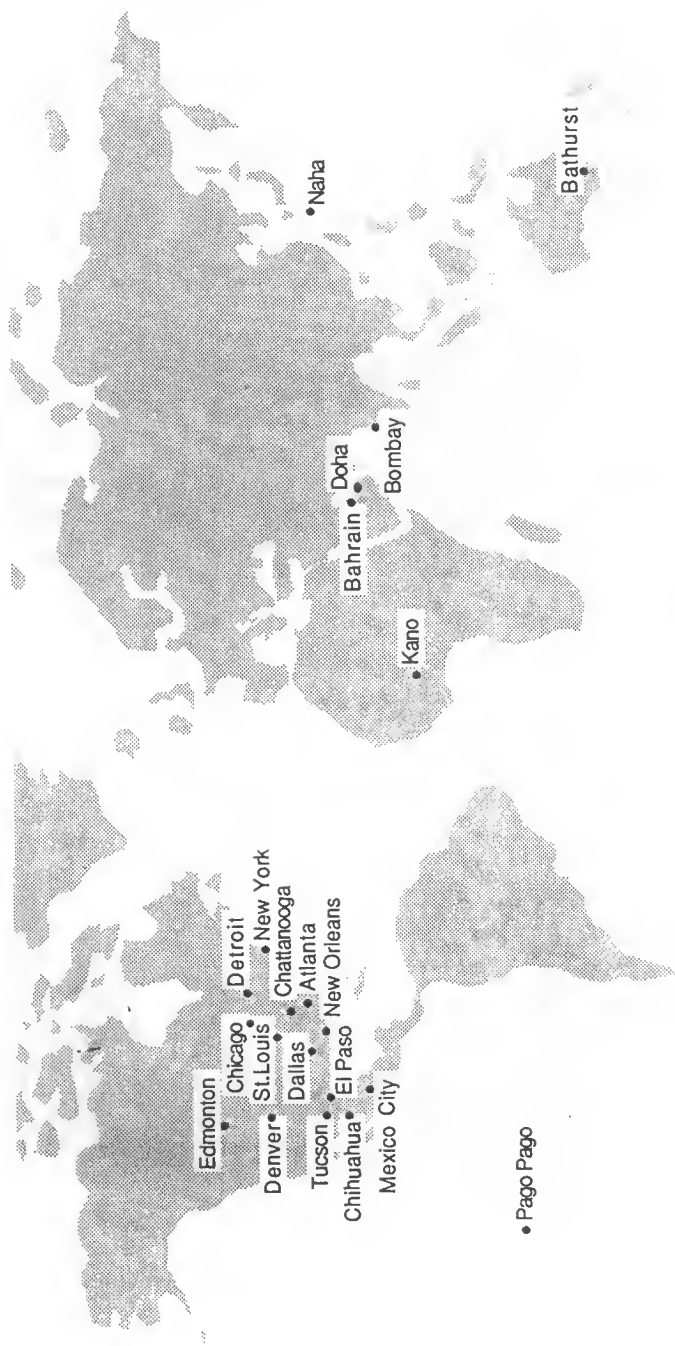


Fig. 1 - Worldwide distribution of aircraft accidents due to windshear.

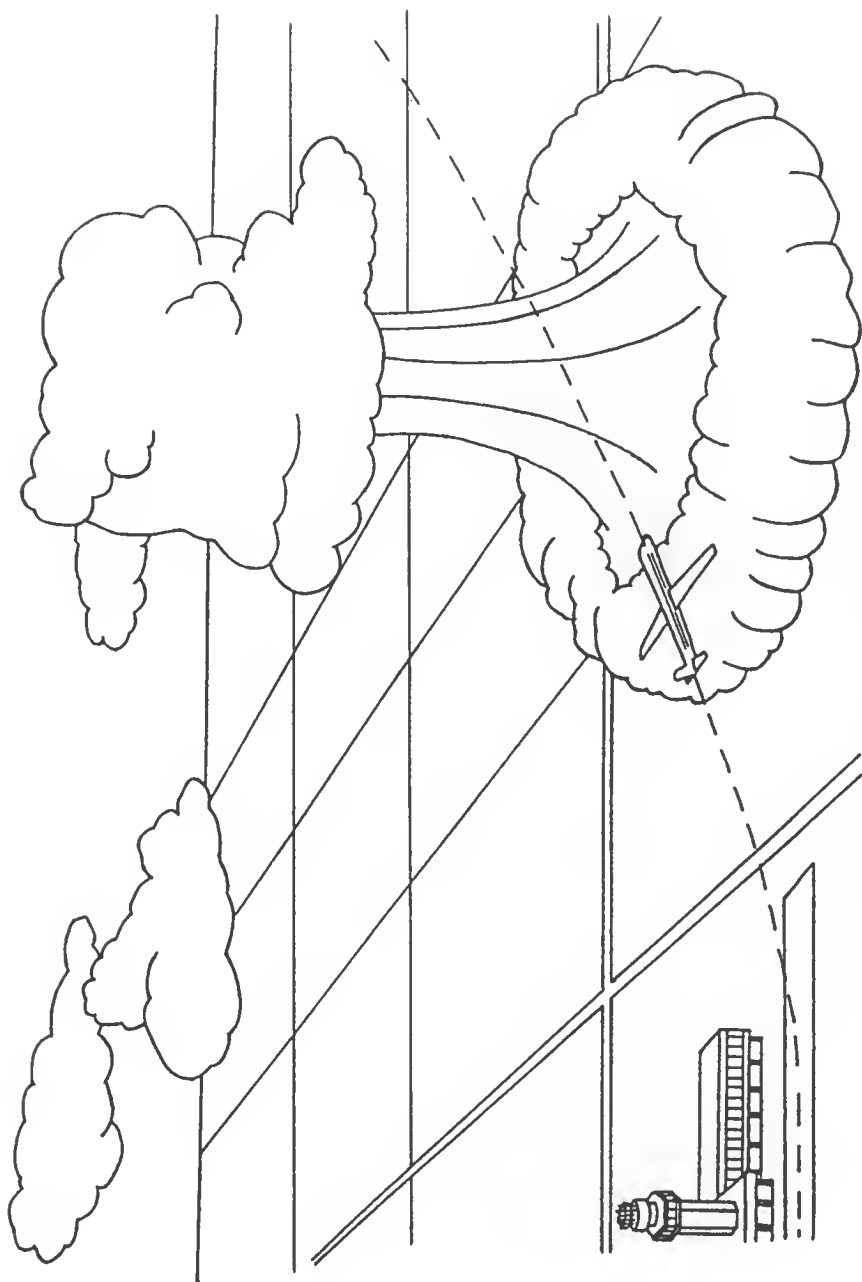


Fig. 2 - Microburst encounter, 3D-view.

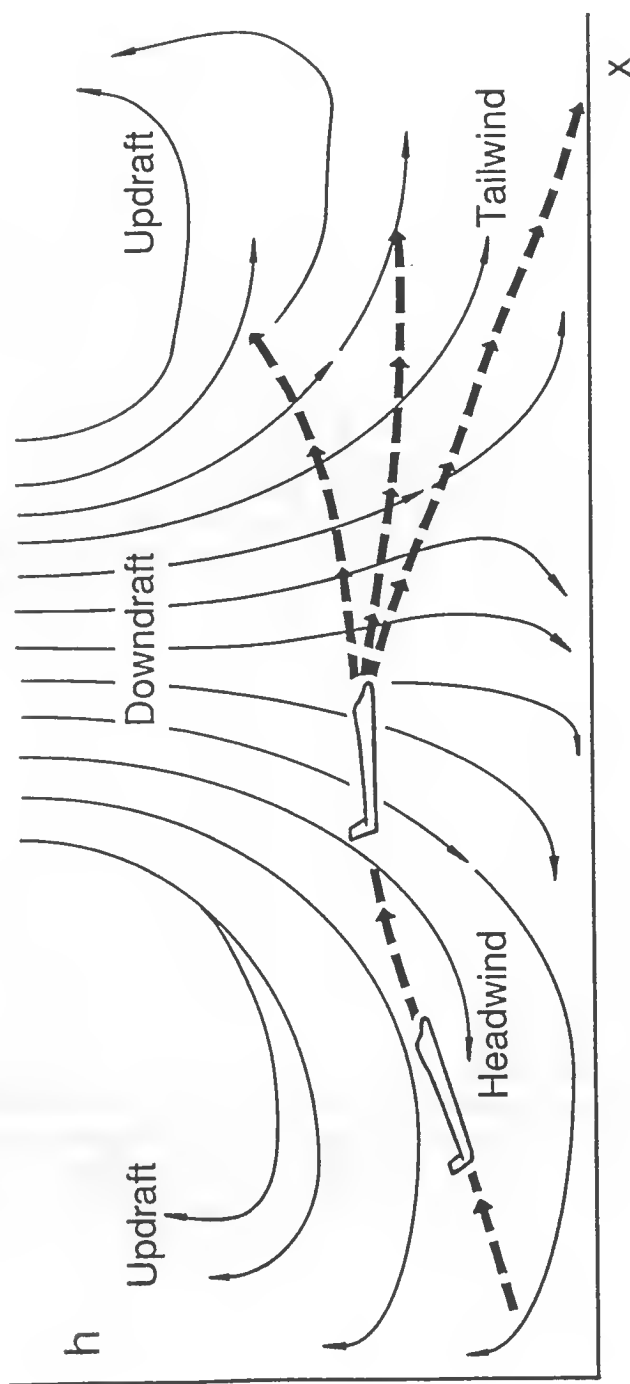


Fig. 3 - Microburst encounter, 2D-view.

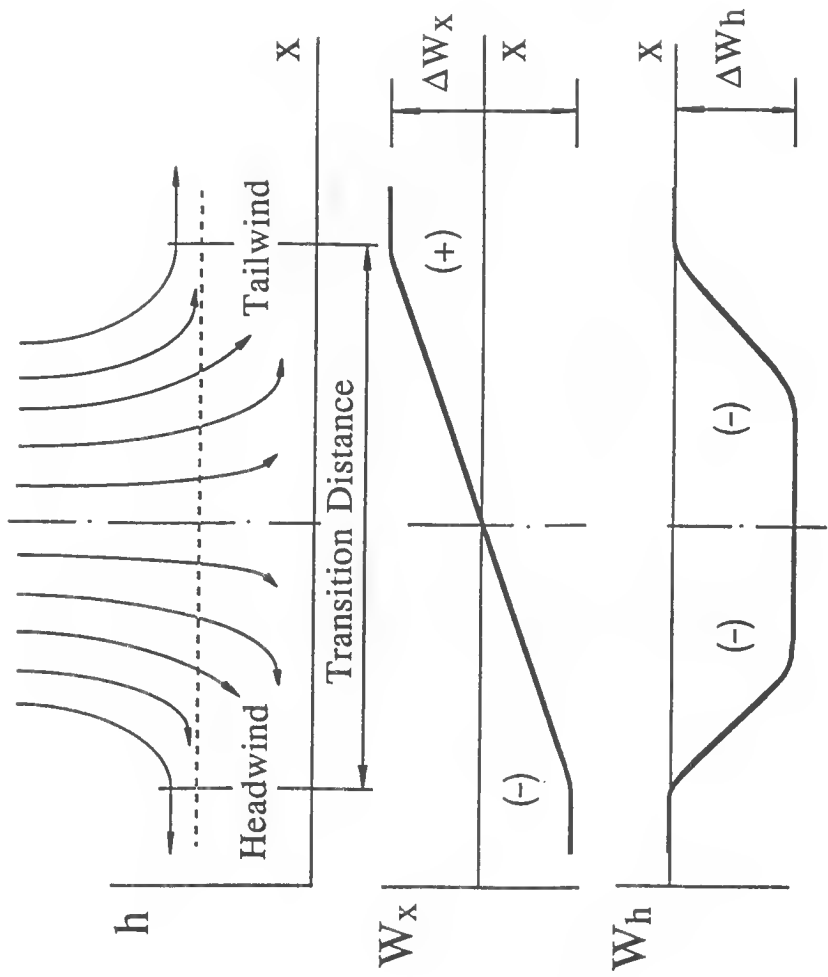


Fig. 4 - Wind velocity components and wind velocity differences.

transition is a transport acceleration, hence a windshear inertia force, which in some cases can be as large as the drag of the aircraft or the thrust of the engines. Under these conditions, the controllability of the aircraft is diminished, occasionally resulting in a crash.

Among the major aircraft accidents of recent years, three are particularly noteworthy: the crash of Eastern Airlines Flight 066 at JFK International Airport (1975); the crash of PANAM Flight 759 at New Orleans International Airport (1982); and the crash of Delta Airlines Flight 191 at Dallas-Fort Worth International Airport (1985). Spurred by these accidents, an intense research program ensued in the United States, with the involvement of government organizations (NASA, FAA), private organizations (ALPA, AREF), industries (Boeing, Lockheed, McDonnell Douglas), airlines (United, American, Delta), and universities (Princeton, Stanford, UCB, Rice).

In this paper, we give an account of the research effort performed at Rice University over the past eight years [Refs. 5-16] on the optimization and guidance of flight trajectories in windshear. We consider three basic problems: take-off, abort landing, and penetration landing.

2. Windshear and Aircraft Accidents

In a nutshell, windshear is synonymous with nonuniformity of the wind: if there is no wind, obviously there is no shear; if there is wind, and the wind is uniform, there is no shear; if there is wind, and the wind is nonuniform, there is shear. The intensity of the shear is measured by the degree of nonuniformity of the wind, hence by the gradients of the wind components. The bigger the gradients, the bigger the shear, hence the greater is the windshear inertia force and the ensuing dynamical effect on the aircraft.

Accidents due to windshear have occurred in both take-off and landing and have involved many different types of aircraft. Some examples are shown in Tables 1-2, where the following information is displayed: flight, airplane type, location, year, number of casualties, and number of injuries. For accidents involving the Boeing 727 aircraft, Tables 3-4 supply windshear data obtained via wind identification procedures.

Table 3 shows the total horizontal wind velocity difference ΔW_x (maximum tailwind minus maximum headwind), the corresponding average gradient $\Delta W_x / \Delta x$, the ratio $\bar{W}_x / g$ of average wind acceleration to acceleration of gravity, and the thrust-to-weight ratio T/W . The ratio

Table 1. Aircraft accidents due to windshear, take-off.

Flight	Airplane	Location	Year	Casualties	Injuries
BOA-252	Argonaut	Kano, Nigeria	1956	32	11
CAL-426	B-727	Denver, Colorado	1975	0	15
CAL-063	B-727	Tucson, Arizona	1977	0	0
PAA-759	B-727	New Orleans, Louisiana	1982	153	9

Table 2. Aircraft accidents due to windshear, landing.

Flight	Airplane	Location	Year	Casualties	Injuries
PAA-806	B-707	Pago Pago, Samoa	1974	96	5
EAL-066	B-727	New York, New York	1975	112	12
RJA-600	DC-9	Doha, Qatar	1976	45	15
ALA-121	DC-9	Philadelphia, Pennsylvania	1976	0	87
DAL-191	L-1011	Dallas, Texas	1985	137	26

$\dot{W}_x/g$ is the same as the ratio of the average windshear inertia force to the aircraft weight. Upon comparing the last two columns, it is clear that, for the accidents reported in Table 3, the average windshear inertia force was of the order of the thrust, albeit smaller than the thrust.

The quantities $\Delta W_x/\Delta x$ and $\dot{W}_x/g$ of Table 3 are averaged over the entire transition length from maximum headwind to maximum tailwind. Larger values of $\Delta W_x/\Delta x$ and $\dot{W}_x/g$ (sustained values) are obtained by computing the averages over smaller distances, albeit distances larger than 1500 ft. This is done in Table 4, which shows that the largest sustained values can be three times the values averaged over the entire transition length. In particular, upon comparing $\dot{W}_x/g$ with T/W , we see

that, for some of the accidents reported in Table 4, the largest sustained windshear inertia force was greater than the thrust.

Tables 3-4 also show supplementary data obtained from a particular windshear episode which took place on August 1, 1983 at Andrews AFB near Washington, DC. In this particular episode, there was no aircraft accident, and the wind velocities were measured by means of anemometers on the ground. The value of ΔW_x recorded at Andrews AFB (about 300 fps) is more than double the largest value recovered from the analysis of aircraft accidents (about 140 fps). It appears that the windshear associated with the Andrews AFB episode could not have been survived by a Boeing 727 aircraft; for that matter, it could not have been survived by Air Force One, which fortunately had landed six minutes earlier with President Reagan on board. However, there is some question as to the accuracy of the Andrews AFB measurements, and it is possible that this question might never be answered.

Table 3. Windshear data, average values, B-727 accidents.

Flight	ΔW_x (fps)	$\Delta W_x/\Delta x$ (1/sec)	$\dot{W}_x/g$	T/W
EAL-066	126	0.010	0.075	0.220
CAL-426	135	0.013	0.114	0.220
CAL-063	126	0.015	0.131	0.220
PAA-759	67	0.008	0.070	0.220
Andrews AFB	294	0.050	0.400	0.220

Table 4. Windshear data, sustained values, B-727 accidents.

Flight	$\Delta W_x/\Delta x$ (1/sec)	$\dot{W}_x/g$	T/W
EAL-066	0.030	0.225	0.220
CAL-426	0.040	0.350	0.220
CAL-063	0.022	0.192	0.220
PAA-759	0.031	0.271	0.220
Andrews AFB	0.060	0.480	0.220

3. Windshear Safety Systems

To offset the windshear threat, there are two basic approaches: avoidance systems and recovery systems. An avoidance system is designed to alert a pilot to the fact that a windshear encounter might take place; here, the intent is the avoidance of a microburst. A recovery system is designed to guide the pilot in the course of a windshear encounter; here, the intent is to fly smartly across a microburst, if an inadvertent encounter takes place. Obviously, avoidance systems and recovery systems are not mutually exclusive, but complementary to one another.

3.1. Avoidance Systems. Among these, we mention the following: ground-based mechanical systems (anemometers), ground-based radar systems (Doppler radar), and airborne systems (radar/lidar and/or aerodynamic/inertial instrumentation).

Ground-based mechanical systems involve the installation of anemometers at different points of an airport. These anemometers provide the wind direction and intensity. If a wind velocity divergence is detected, and if the divergence exceeds a certain threshold, an alert is issued. Depending on the situation, the director of an airport can choose among several options: he can alert the pilots of planes flying nearby; he can shut down a particular runway; or he can shut down the entire airport. An example is the low-level windshear alert system (LLWAS); this system is simple, has low cost, but does not have great accuracy.

Ground-based radar systems involve the installation of Doppler radars, capable of measuring the wind distribution in the space around an airport. The difficulty of this technique is the presence of ground clutter, which can have a significant effect on the signal-to-noise ratio. Compared with anemometer systems, ground-based radar systems are more accurate, but their cost is high.

Airborne systems involve the installation on an airplane of either a radar or a lidar, capable of seeing in the direction along which the airplane is flying. Compared with ground-based radar systems, airborne systems have a lower signal-to-noise ratio because of limitations on transmitter power and antenna/telescope size. Airborne systems have less accuracy than ground-based radar systems.

Alternative airborne systems involve the combination of aerodynamic and inertial instrumentation. The aerodynamic instrumentation supplies the airspeed, hence the velocity of the aircraft relative to the air mass, hence the relative acceleration (derivative of the relative velocity). The inertial instrumentation supplies the absolute acceleration. The

difference between the absolute acceleration and the relative acceleration is the wind acceleration, which in turn is related to the windshear. Whenever the wind acceleration exceeds a certain threshold, an alert is issued.

3.2. Recovery Systems. Among these, we mention the following: maximum angle of attack guidance, constant pitch guidance, and advanced guidance.

In the maximum angle of attack guidance, the aircraft is rotated at the maximum permissible rate (3 deg/sec) until the stick-shaker angle of attack is reached; afterward, the aircraft is held at this angle of attack. This system has been in use for some 20 years and does not require any new instrumentation. However, its performance is poor in both take-off and abort landing, as shown by computer simulations.

In the constant pitch guidance, the aircraft is rotated at the maximum permissible rate until a certain target pitch angle is obtained (15 deg of fuselage pitch). Afterward, the aircraft is maintained at the target pitch angle, subject to the stick-shaker angle-of-attack limitation. This system was adopted in the 1987 Windshear Training Aid [Ref. 2] and does not require any new instrumentation. Its performance is better than that of the maximum angle of attack guidance, but not as good as that of an advanced guidance system.

Advanced guidance systems are derived from the properties of the optimal trajectories and are constructed in such a way that they approximate the behavior of the optimal trajectories while using only local/prior information on the windshear and the downdraft. There are two possible operational modes: the totally automated mode and the semiautomated or display mode; the latter is preferred for flight operations.

In the semiautomated mode, an advanced guidance system includes three parts: sensors, software, and display. The sensors supply the local measurements of the windshear, the downdraft, and certain components of the state of the aircraft. The software consists of a feedback-control algorithm, which processes the information supplied by the sensors and arrives at the guidance values for either the pitch angle or the angle of attack. These values lead toward the satisfaction of a guidance law, obtained from the analysis of the optimal trajectories. The display shows either the actual pitch and the guidance pitch or the actual angle of attack and the guidance angle of attack. Then, in the semiautomated mode, the pilot must act on the controls in such a way that the actual pitch approaches the guidance pitch or the actual angle of attack approaches the guidance angle of attack.

An advanced guidance system requires the installation of instrumentation measuring the wind acceleration and the downdraft. These quantities can be computed from measurements already available in aircraft equipped with aerodynamic/inertial instrumentation, such as the Boeing B-747 and the Lockheed L-1011. On the other hand, for older-type aircraft (for instance, the Boeing B-727), the installation of additional sensors is necessary. The implementation of this system is not too costly.

Examples of advanced guidance systems are the acceleration guidance, the gamma guidance, and the penetration landing guidance, developed by the Aero-Astronautics Group of Rice University.

4. Wind Model

Basically, the downburst can be assumed to be an axysymmetric phenomenon [Ref. 17]. In its simplest form, the flow field can be modelled by postulating the presence of two horizontal vortex rings of radius R [Ref. 18]: the primary ring, having circulation $+\Gamma$, is located at altitude $+H$ above ground; the image ring, having circulation $-\Gamma$, is located at altitude $-H$ below ground. If one assumes steady flow and uses either the Biot-Savart law or the properties of the stream function, one arrives at a wind velocity distribution of the form ⁽¹⁾

$$(1) \quad W_x = \lambda A(x, h), \quad W_h = \lambda B(x, h) .$$

Here, x is a radial coordinate; h is a vertical coordinate; and λ is a parameter proportional to the circulation strength Γ . While the functions $A(x, h)$ and $B(x, h)$ depend on the parameters H and R , they are independent of λ .

Suppose that, at each given altitude, one computes the horizontal wind velocity difference ΔW_x (maximum tailwind minus maximum headwind) and the vertical wind velocity difference ΔW_h (maximum updraft minus maximum downdraft). Because the above maxima and minima are computed with respect to x , the result is that the wind velocity

⁽¹⁾ Sign convention for wind velocity components: $W_x > 0$ for tailwind and $W_x < 0$ for headwind; also, $W_h > 0$ for updraft and $W_h < 0$ for downdraft.

differences depend only on the altitude; hence, they take the form

$$(2) \quad \Delta W_x = \lambda F(h), \quad \Delta W_h = \lambda G(h) .$$

4.1. Low Altitude Approximation. For altitudes near the ground, specifically for $0 \leq h \leq 1000$ ft, the function $A(x, h)$ is nearly independent of the altitude, while the function $B(x, h)$ is nearly linear with the altitude [Ref. 19]. This is the same as stating that Eqs. (1) can be replaced by ⁽²⁾

$$(3) \quad W_x = \lambda A(x), \quad W_h = \lambda (h/h_*) B(x) ,$$

while Eqs. (2) can be replaced by

$$(4) \quad \Delta W_x = \lambda \Delta W_{x*}, \quad \Delta W_h = \lambda (h/h_*) \Delta W_{h*} .$$

Here, $h_* = 1000$ ft is a reference altitude; $\Delta W_{x*} = 100$ fps is a reference value of the horizontal wind velocity difference ($\lambda = 1$); and ΔW_{h*} is a reference value of the vertical wind velocity difference ($\lambda = 1$ and $h = h_*$).

Decreasing values of λ (hence, smaller ΔW_x) correspond to milder windshears; conversely, increasing values of λ (hence, greater ΔW_x) correspond to more severe windshears. Therefore, by changing the value of λ , one can generate shear/downdraft combinations ranging from extremely mild to extremely severe.

Figure 5 shows the functions $W_x(x)$ and $W_h(x, h)$, computed for $\lambda = 1$, corresponding to $\Delta W_x = 100$ fps. For other values of λ (hence, for other values of ΔW_x), the ordinates in Fig. 5 must be scaled proportionally to λ (hence, proportionally to ΔW_x).

4.2. Comment. Although no two windshear encounters are exactly alike, two basic phenomena are always present: shear and downdraft. Therefore, it is important that these essential characteristics be present in the wind model employed in optimization and guidance studies. This is precisely the case with the wind model represented by Eqs. (3)-(4) and Fig. 5, which has the following properties: (a) there is a transition from

⁽²⁾ Although the same notation is employed in Eqs. (1)-(2) and Eqs. (3)-(4), the symbols A , B , λ have different meaning.

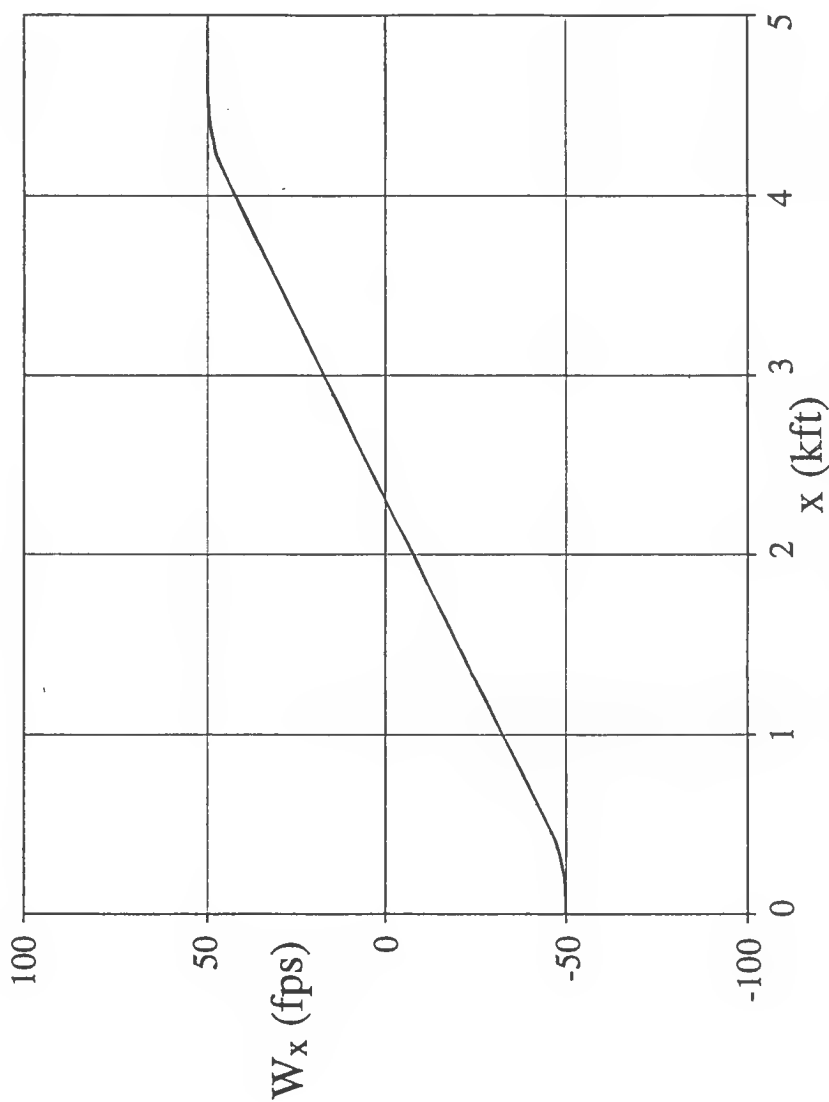


Fig. 5A - Horizontal wind vs distance.

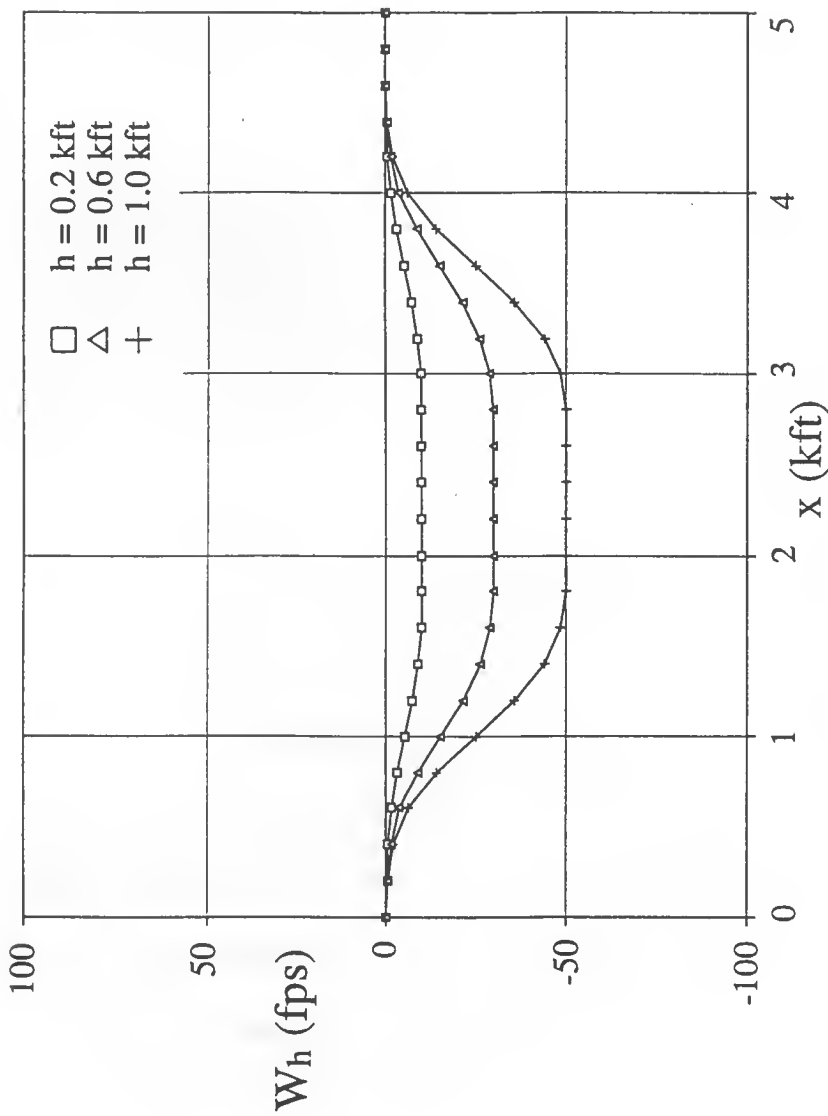


Fig. 5B - Vertical wind vs distance.

a uniform headwind to a uniform tailwind, with nearly constant shear in the core of the downburst; (b) the downdraft achieves maximum negative value at the center of the downburst; (c) the downdraft vanishes on the ground, $h=0$; and (d) the wind velocity components nearly satisfy the continuity equation and the irrotationality condition in the core of the downburst.

5. Equations of Motion

In this paper, we make use of the relative wind-axes system in connection with the following assumptions: (a) the aircraft is a particle of constant mass; (b) flight takes place in a vertical plane; (c) Newton's law is valid in an Earth-fixed system; and (d) the wind flow field is steady.

With the above premises, the equations of motion include the kinematical equations

$$(5a) \quad \dot{x} = V \cos \gamma + W_x ,$$

$$(5b) \quad \dot{h} = V \sin \gamma + W_h ,$$

and the dynamical equations

$$(6a) \quad \dot{V} = (T/m) \cos(\alpha + \delta) - D/m - g \sin \gamma - (\dot{W}_x \cos \gamma + \dot{W}_h \sin \gamma),$$

$$(6b) \quad \dot{\gamma} = (T/mV) \sin(\alpha + \delta) + L/mV - (g/V) \cos \gamma + (1/V) (\dot{W}_x \sin \gamma - \dot{W}_h \cos \gamma) .$$

Because of assumption (d), the total derivatives of the wind velocity components and the corresponding partial derivatives satisfy the relations

$$(7a) \quad \dot{W}_x = (\partial W_x / \partial x) (V \cos \gamma + W_x) + (\partial W_x / \partial h) (V \sin \gamma + W_h) ,$$

$$(7b) \quad \dot{W}_h = (\partial W_h / \partial x) (V \cos \gamma + W_x) + (\partial W_h / \partial h) (V \sin \gamma + W_h) .$$

These equations must be supplemented by the functional relations

$$(8a) \quad T = T(h, V, \beta) ,$$

$$(8b) \quad D = D(h, V, \alpha), \quad L = L(h, V, \alpha) ,$$

$$(8c) \quad W_x = W_x(x, h), \quad W_h = W_h(x, h),$$

and by the analytical relations

$$(9a) \quad \theta = \alpha + \gamma,$$

$$(9b) \quad \gamma_e = \arctan [(V \sin \gamma + W_h)/(V \cos \gamma + W_x)] .$$

In the light of (7)-(8), the differential system (5)-(6) involves four state variables $[x(t), h(t), V(t), \gamma(t)]$ and two control variables $[\alpha(t), \beta(t)]$. However, the number of control variables reduces to one (the angle of attack), if the power setting is specified in advance. The quantities (9) can be computed a posteriori, once the values of the state and the control are known.

5.1. Shear/Downdraft Factor. Let E denote the aircraft energy per unit weight,

$$(10) \quad E = h + V^2/2g .$$

Upon taking the time derivative of both sides of Eq. (10), and upon accounting for Eqs. (5b) and (6a), we obtain the relation

$$(11) \quad \dot{E} = [(T/W) \cos(\alpha + \delta) - D/W - F] V .$$

Here, F is the shear/downdraft factor [Ref. 10],

$$(12a) \quad F = (\dot{W}_x/g) \cos \gamma + (\dot{W}_h/g) \sin \gamma - W_h/V ,$$

which simplifies to

$$(12b) \quad F \cong \dot{W}_x/g - W_h/V ,$$

if γ is small and if $\dot{W}_h$ is at most of the same order as $\dot{W}_x$. Clearly, the shear/downdraft factor combines the effects of the shear and the downdraft into a single entity. Note that $F > 0$ for an unfavorable shear, $F < 0$ for a favorable shear, and $F = 0$ if there is no shear nor downdraft. The product FV represents the change of the energy time rate due to the combined effect of the shear and the downdraft. This is why the quantity defined by either of Eqs. (12) has considerable importance in guidance.

5.2. Inequality Constraints. The angle of attack α and its time derivative $\dot{\alpha}$ are subject to the inequalities

$$(13a) \quad \alpha \leq \alpha_*,$$

$$(13b) \quad -\dot{\alpha}_* \leq \dot{\alpha} \leq \dot{\alpha}_*,$$

where α_* is a prescribed upper bound (the stick-shaker angle of attack) and $\dot{\alpha}_*$ is a positive constant (e.g., $\dot{\alpha}_* = 3.0$ deg/sec).

The power setting β and its time derivative $\dot{\beta}$ are subject to the inequalities

$$(14a) \quad \beta_* \leq \beta \leq 1,$$

$$(14b) \quad -\dot{\beta}_* \leq \dot{\beta} \leq \dot{\beta}_*,$$

where β_* is a prescribed lower bound (e.g., $\beta_* = 0.2$) and $\dot{\beta}_*$ is a positive constant (e.g., $\dot{\beta}_* = 0.2/\text{sec}$).

6. Take-Off Optimization

In take-off under adverse weather conditions, a prudent pilot employs maximum power setting,

$$(15) \quad \beta = 1, \quad 0 \leq t \leq \tau.$$

Hence, the only remaining control is the angle of attack $\alpha(t)$, which is subject to Ineqs. (13).

Optimal take-off trajectories can be determined by minimizing the peak value of the deviation of the absolute path inclination from a reference value. Thus, the performance index has the form

$$(16a) \quad I_* = \min_{\alpha} I,$$

$$(16b) \quad I = \max_t |\gamma_e - \gamma_{eR}|, \quad 0 \leq t \leq \tau,$$

$$(16c) \quad \gamma_e = \arctan [(V \sin \gamma + W_h)/(V \cos \gamma + W_x)].$$

Here, α is the angle of attack; γ_e is the absolute path inclination;

$\gamma_{eR} = \gamma_{e0}$ is a reference value for γ_e ; γ is the relative path inclination; V is the relative velocity; and W_x , W_h are the wind components.

This is a minimax problem or Chebyshev problem of optimal control, which can be converted into a Bolza problem via suitable transformations. Specifically, the performance index (16) is replaced by

$$(17a) \quad J_* = \min_{\alpha} J ,$$

$$(17b) \quad J = \int_0^{\tau} (\gamma_e - \gamma_{eR})^q dt ,$$

$$(17c) \quad \gamma_e = \arctan [(V \sin \gamma + W_h) / (V \cos \gamma + W_x)] ,$$

where q is some large, positive, even integer (for instance, $q=6$).

The Bolza problem represented by the performance index (17), the differential and inequality constraints of Section 5, and suitable boundary conditions is amenable to computational solution by means of the sequential gradient-restoration algorithm; see Ref. 20 for a bibliography and recent versions of this algorithm.

6.1. Data and Results. The aircraft under consideration is the Boeing B-727 with take-off weight $W=180000$ lb, gear up, flap deflection $\delta_F=15.0$ deg, and stick-shaker angle of attack $\alpha_* = 16.0$ deg. At the initial point, the state variables are given (in particular, $h_0=50$ ft). At the final point, gamma recovery is required ($\gamma_\tau = \gamma_0$), where $\tau=40$ sec is the prescribed final time.

Optimal trajectories were computed for the above aircraft and the wind model of Section 4. Three values of the wind intensity parameter were employed: $\lambda=0.8, 1.0, 1.1$. These values correspond to the following wind velocity differences: $\Delta W_x=80, 100, 110$ fps.

Figure 6 shows the altitude profile $h(t)$, relative velocity profile $V(t)$, and angle of attack profile $\alpha(t)$. Clearly, as ΔW_x changes, the altitude profile of the optimal trajectory changes: it is entirely ascending for $\Delta W_x=80$ fps; it includes a nearly-horizontal branch for $\Delta W_x=100$ fps; and it includes a descending branch for $\Delta W_x=110$ fps. Should the wind velocity difference be further increased to $\Delta W_x=120$ fps, the B-727 would crash, even flying an optimal take-off trajectory.

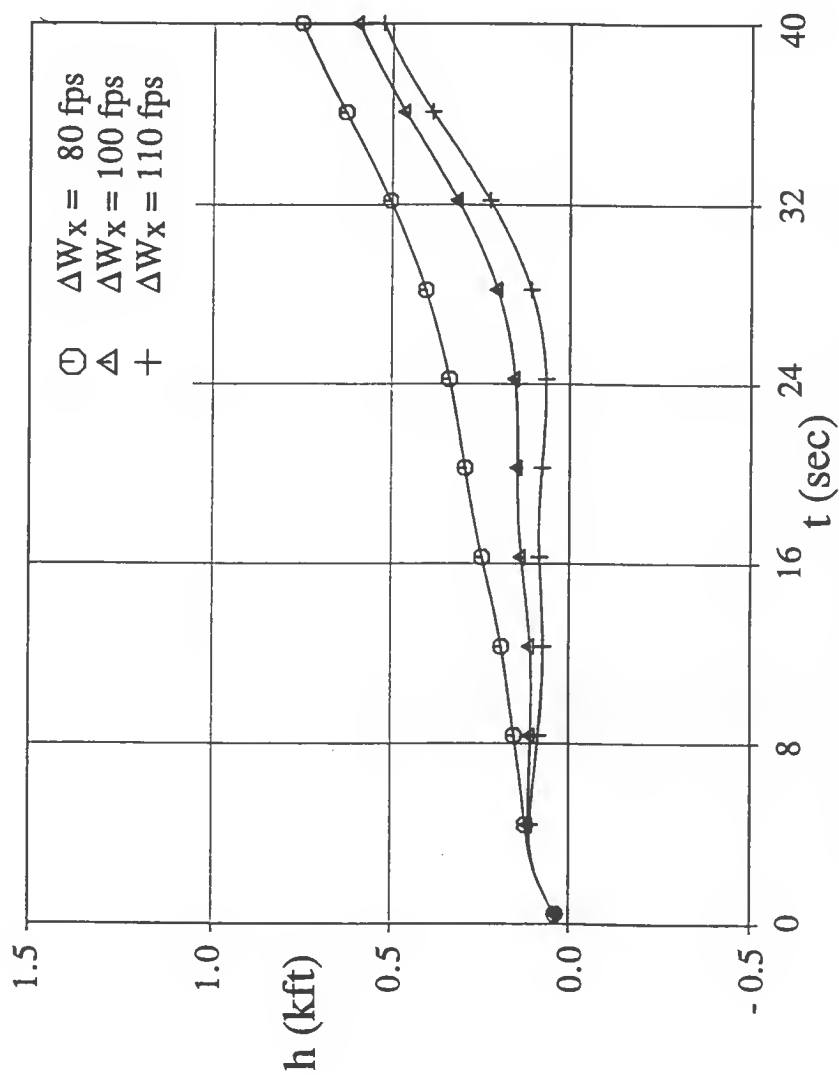


Fig. 6A - Optimal take-off trajectories, altitude vs time.

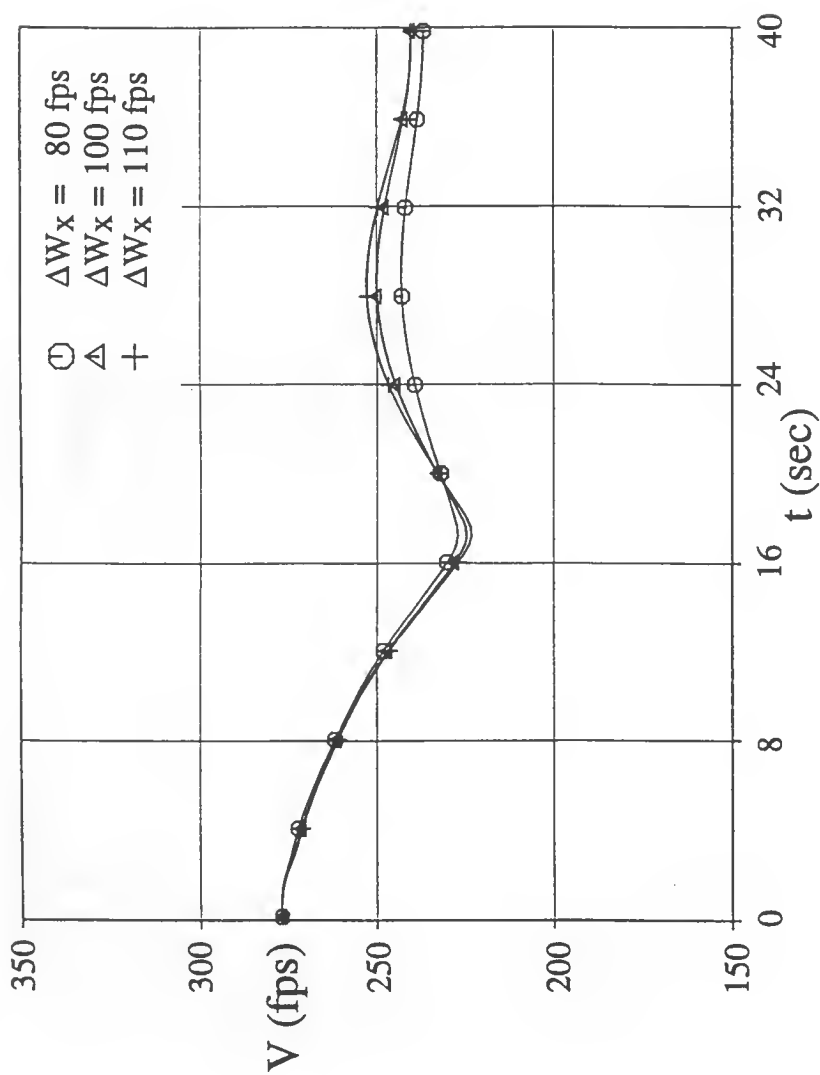


Fig. 6B - Optimal take-off trajectories, relative velocity vs time.

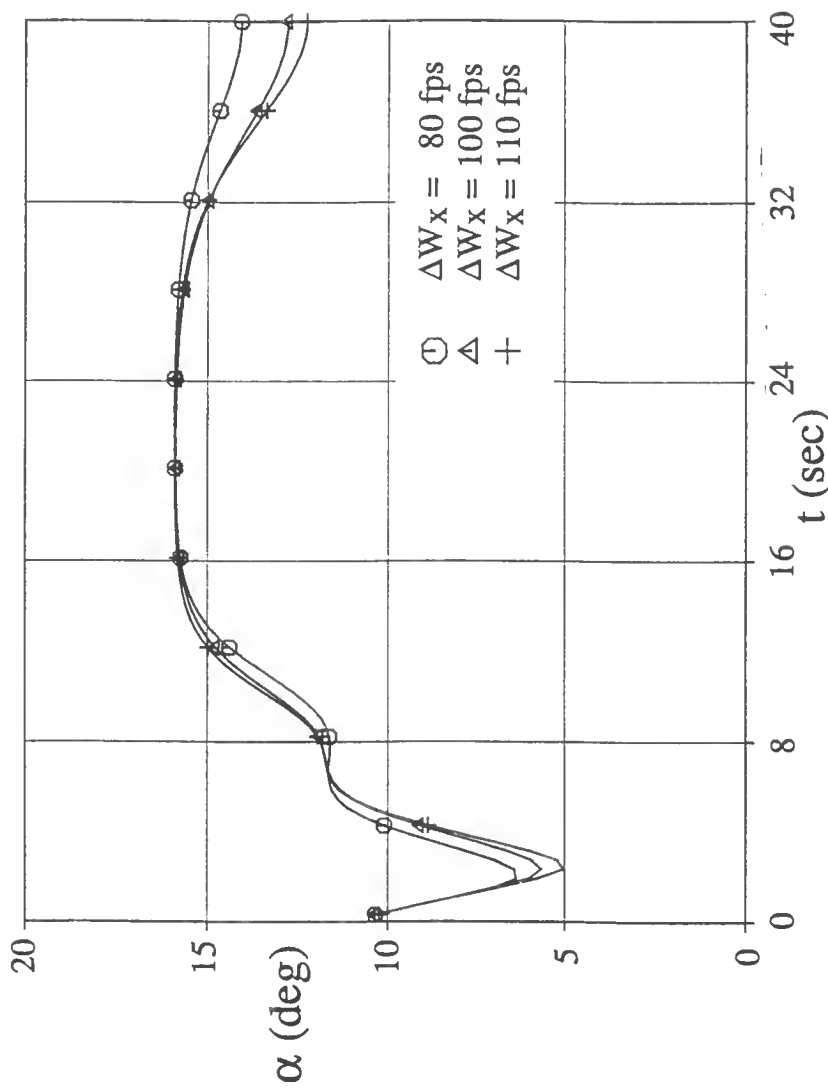


Fig. 6C - Optimal take-off trajectories, angle of attack vs time.

6.2. Extensions. The analyses done for the B-727 aircraft were extended to the B-737 and B-747 aircraft, leading to the following general conclusions [Ref. 21]:

- (i) the optimal trajectories achieve minimum velocity at the end of the shear;
- (ii) the optimal trajectories require an initial decrease in the angle of attack, followed by a gradual increase; the stick-shaker angle of attack is achieved near the end of the shear;
- (iii) for weak-to-moderate windshears, the optimal trajectories are characterized by continuous climb; the average value of the path inclination decreases as the intensity of the shear increases;
- (iv) for relatively severe windshears, the optimal trajectories are characterized by an initial climb, followed by nearly-horizontal flight, followed by renewed climbing after the aircraft has passed through the shear region;
- (v) extremely severe windshears are not survivable, even flying an optimal take-off trajectory.

7. Take-Off Guidance

The computation of the optimal trajectories requires global information on the wind field; that is, it requires the knowledge of the wind components in the region of space where the aircraft is flying. In practice, global information is not available; even if it were, there would not be enough computing capability onboard and enough time to process it adequately. As a consequence, the optimal trajectories are merely a benchmark that one should strive to approach in flight.

An alternative to using global information is to employ local/prior information on the wind acceleration and the downdraft. Therefore, the guidance problem must be addressed in these terms: assuming that local/prior information is available on the wind acceleration, the downdraft, as well as certain components of the state of the aircraft, we wish to guide an aircraft automatically or semiautomatically in such a way that the key properties of the optimal trajectories are preserved.

Based on the idea of preserving the key properties of the optimal trajectories, two guidance schemes were developed at Rice University: (AG) acceleration guidance, based on the relative acceleration; (GG) gamma guidance, based on the absolute path inclination.

7.1. Acceleration Guidance. Let (R1) and (R2) denote the unfavorable shear region and the aftershear region, respectively. In feedback control form, the acceleration guidance (AG) is represented by

$$(18a) \quad \alpha - \bar{\alpha}(V) = K_1 (\dot{V}/g + C_1 F), \quad \text{Region (R1),}$$

$$(18b) \quad \alpha - \bar{\alpha}(V) = K_2 (V - C_2 V_0), \quad \text{Region (R2).}$$

In Eqs. (18), K_1 and K_2 are the gain coefficients (positive); C_1 and C_2 are positive constants; α is the guidance angle of attack; $\bar{\alpha}(V)$ is the nominal angle of attack associated with static equilibrium in the direction normal to the flight path in the absence of shear and downdraft; V is the relative velocity; V_0 is the initial relative velocity; and F is the shear/downdraft factor (12). The meaning of Eqs. (18) is the following: in Region (R1), the deceleration increases as the shear/downdraft factor increases; in Region (R2), the relative velocity tends to some prescribed fraction of the initial value.

7.2. Gamma Guidance. In feedback control form, the gamma guidance (GG) is represented by

$$(19a) \quad \alpha - \bar{\alpha}(V) = -K_1 [\gamma_e - \gamma_{e0} (1 - C_1 F)], \quad \text{Region (R1),}$$

$$(19b) \quad \alpha - \bar{\alpha}(V) = -K_2 (\gamma_e - C_2 \gamma_{e0}), \quad \text{Region (R2).}$$

In Eqs. (19), K_1 and K_2 are the gain coefficients (positive); C_1 and C_2 are positive constants; α is the guidance angle of attack; $\bar{\alpha}(V)$ is the nominal angle of attack; γ_e is the absolute path inclination; γ_{e0} is the initial absolute path inclination; and F is the shear/downdraft factor (12). The meaning of Eqs. (19) is the following: in Region (R1), the absolute path inclination decreases as the shear/downdraft factor increases; in Region (R2), the absolute path inclination tends to some fraction of the initial value.

7.3. Remark. In both the acceleration guidance and the gamma guidance, suitable switch conditions govern the transition from Region (R1) to Region (R2). For details, see Refs. 10 and 12. Also, for engineering limitations to (18) and (19), see Refs. 10 and 12.

7.4. Trajectory Comparison. Using the above schemes, guidance trajectories were computed for the Boeing B-727 aircraft. It was found that

both the acceleration guidance (AG) and the gamma guidance (GG) trajectories are close to the optimal trajectories (OT). In addition, the resulting near-optimal trajectories are superior to those arising from maximum angle of attack guidance (MAAG) and constant pitch guidance (CPG). See Fig. 7, which refers to $\lambda = 1.0$, hence $\Delta W_x = 100$ fps. While the MAAG trajectory crashes, the CPG trajectory scrapes the ground. On the other hand, the AG trajectory is close to the OT, which has good altitude margin. An analogous remark holds for the GG trajectory, which is not shown, since it is nearly identical to the AG trajectory.

7.5. Survival Capability. The merit of a guidance scheme can be assessed using the concept of survival capability. Consider the one-parameter family of wind models (3)-(4). As λ increases, more intense shear/downdraft combinations are generated until a critical value λ_c (hence, a critical value ΔW_{xc}) is found, such that the minimum altitude equals the ground level.

The results are shown in Table 5, which supplies the survival capability ΔW_{xc} for the optimal trajectory and various guidance trajectories. Table 5 also shows the windshear efficiency ratio WER, defined to be the ratio of the survival capability of a particular trajectory to that of the optimal trajectory. Clearly, the windshear efficiency of the OT is 100%, that of the AG trajectory is 95%, that of the GG trajectory is 96%, that of the CPG trajectory is 85%, and that of the MAAG trajectory is 48%.

Table 5. Survival capability, take-off, B-727.

Trajectory	h_0 (ft)	ΔW_{xc} (fps)	WER
OT	50	119.5	1.000
AG	50	113.0	0.946
GG	50	115.3	0.965
CPG	50	101.8	0.852
MAAG	50	57.7	0.483

Transition distance = 4000 ft

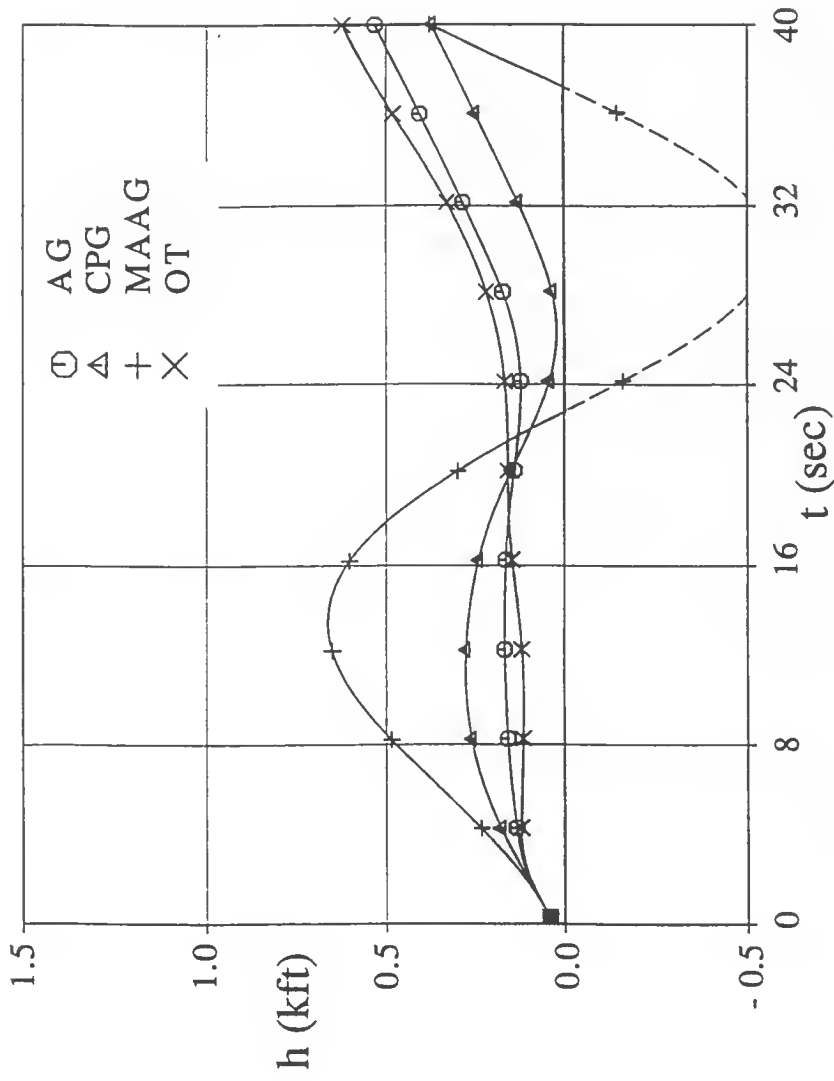


Fig. 7 - Comparison of take-off trajectories, altitude vs time, $\Delta W_x = 100$ fps.

8. Abort Landing Optimization

For an aircraft on an approach path under adverse weather conditions, the pilot has two options: abort landing and penetration landing. The first of these options is analyzed in Sections 8-9; the second is analyzed in Sections 10-11.

In the abort landing maneuver, a prudent pilot performs a quick transition from the approach power setting to the maximum power setting; hence,

$$(20a) \quad \beta = \beta_0 + \dot{\beta}_* t, \quad 0 \leq t \leq \sigma,$$

$$(20b) \quad \beta = 1, \quad \sigma \leq t \leq \tau.$$

In Eqs. (20), $\sigma = (1 - \beta_0)/\dot{\beta}_*$ is the time at which maximum power setting is reached, τ is the final time, and $\dot{\beta}_*$ is the maximum power setting time rate. With β saturated, the only remaining control is the angle of attack $\alpha(t)$, which is subject to Ineqs. (13).

Optimal abort landing trajectories can be determined by minimizing the peak value of the altitude drop. Thus, the performance index has the form

$$(21a) \quad I_* = \min_{\alpha} I,$$

$$(21b) \quad I = \max_t |h - h_R|, \quad 0 \leq t \leq \tau.$$

Here, α is the angle of attack, h is the altitude, and h_R is a reference altitude⁽³⁾.

This is a minimax problem or Chebyshev problem of optimal control, which can be converted into a Bolza problem via suitable transformations. Specifically, the performance index (21) is replaced by

$$(22a) \quad J_* = \min_{\alpha} J,$$

$$(22b) \quad J = \int_0^{\tau} (h - h_R)^q dt,$$

where q is some large, positive, even integer (for instance, $q = 6$).

⁽³⁾ The reference altitude must be such that $h_R - h \geq 0$ in the time interval $0 \leq t \leq \tau$.

The Bolza problem represented by the performance index (22), the differential and inequality constraints of Section 5, and suitable boundary conditions is amenable to computational solution by means of the sequential gradient-restoration algorithm; see Ref. 20 for a bibliography and recent versions of this algorithm.

8.1. Data and Results. The aircraft under consideration is the Boeing B-727 with landing weight $W=150000$ lb, gear down, flap deflection $\delta_F=30.0$ deg, and stick-shaker angle of attack $\alpha_*=17.2$ deg. At the initial point, the state variables are given (in particular, $h_0=600$ ft and $\gamma_{e0}=-3.0$ deg). At the final point, gamma recovery is required ($\gamma_\tau=\gamma_*$), where $\tau=40$ sec is the prescribed final time and γ_* is the relative path inclination for steepest climb.

Optimal trajectories were computed for the above aircraft and the wind model of Section 4. Three values of the wind intensity parameter were employed: $\lambda=1.0, 1.2, 1.4$. These values correspond to the following wind velocity differences: $\Delta W_x=100, 120, 140$ fps.

Figure 8 shows the altitude profile $h(t)$, relative velocity profile $V(t)$, and angle of attack profile $\alpha(t)$. We see that, as ΔW_x increases, the minimum altitude of the optimal trajectory decreases, although the aircraft still survives the windshear encounter. Should the wind velocity difference be further increased to $\Delta W_x=190$ fps, the B-727 would crash, even flying an optimal abort landing trajectory.

8.2. Extensions. The analyses done for the B-727 aircraft were extended to the B-737 and B-747 aircraft, leading to the following general conclusions [Ref. 21]:

- (i) the optimal trajectories include three branches: a descending flight branch, followed by a nearly-horizontal flight branch, followed by an ascending flight branch after the aircraft has passed through the shear region;
- (ii) the optimal trajectories achieve minimum velocity at the end of the shear;
- (iii) the optimal trajectories require an initial decrease of the angle of attack, followed by a gradual increase; the stick-shaker angle of attack is achieved at the end of the shear;
- (iv) the peak altitude drop increases as the windshear intensity increases and the initial altitude increases;
- (v) extremely severe windshears are not survivable, even flying an optimal abort landing trajectory.

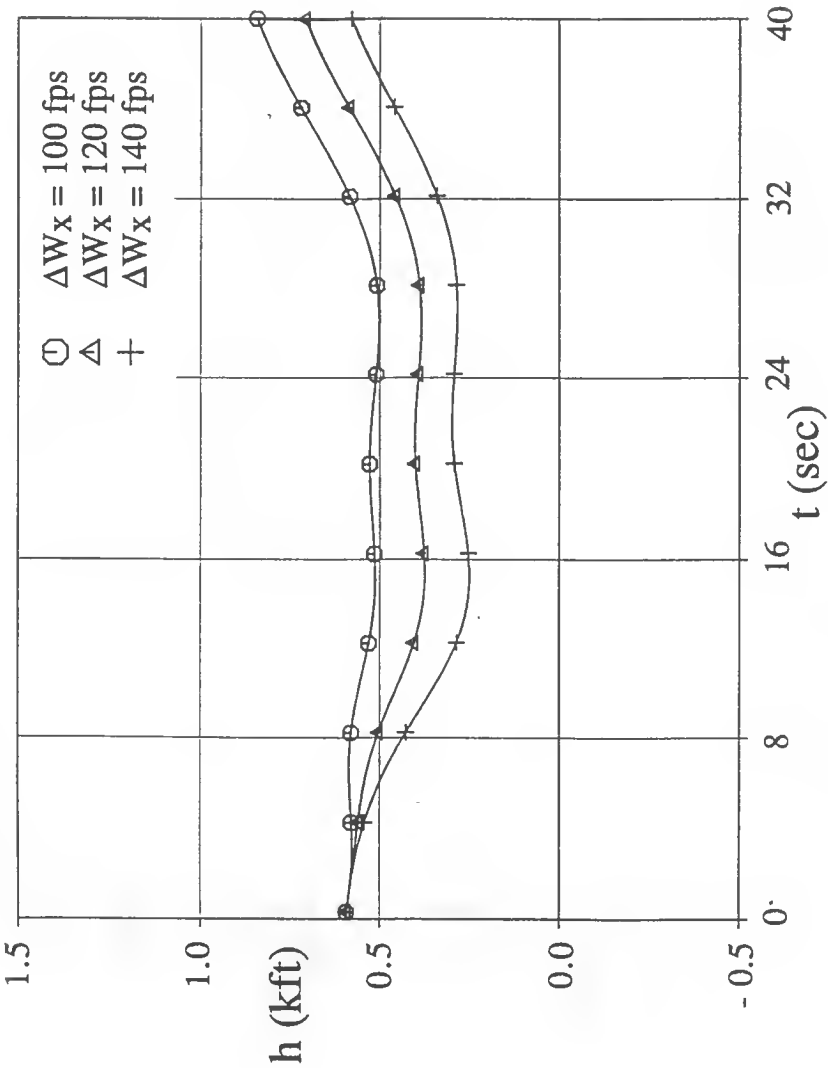


Fig. 8A - Optimal abort landing trajectories, altitude vs time.

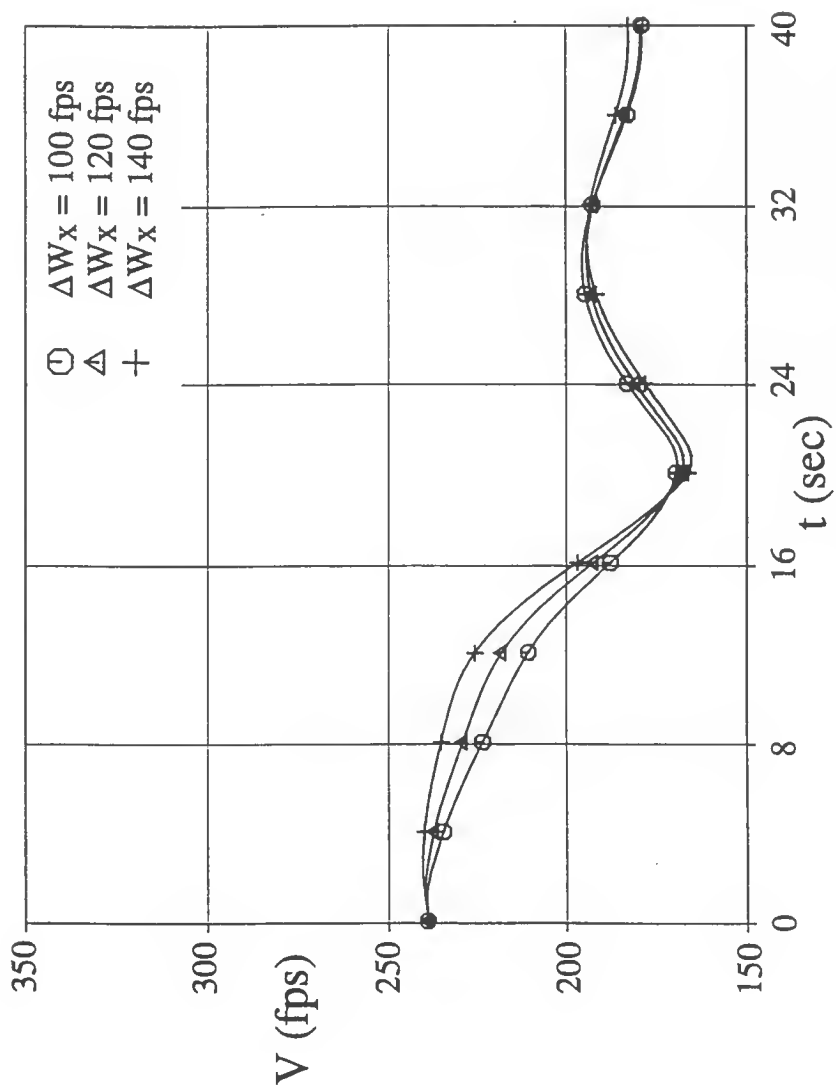


Fig. 8B - Optimal abort landing trajectories, relative velocity vs time.

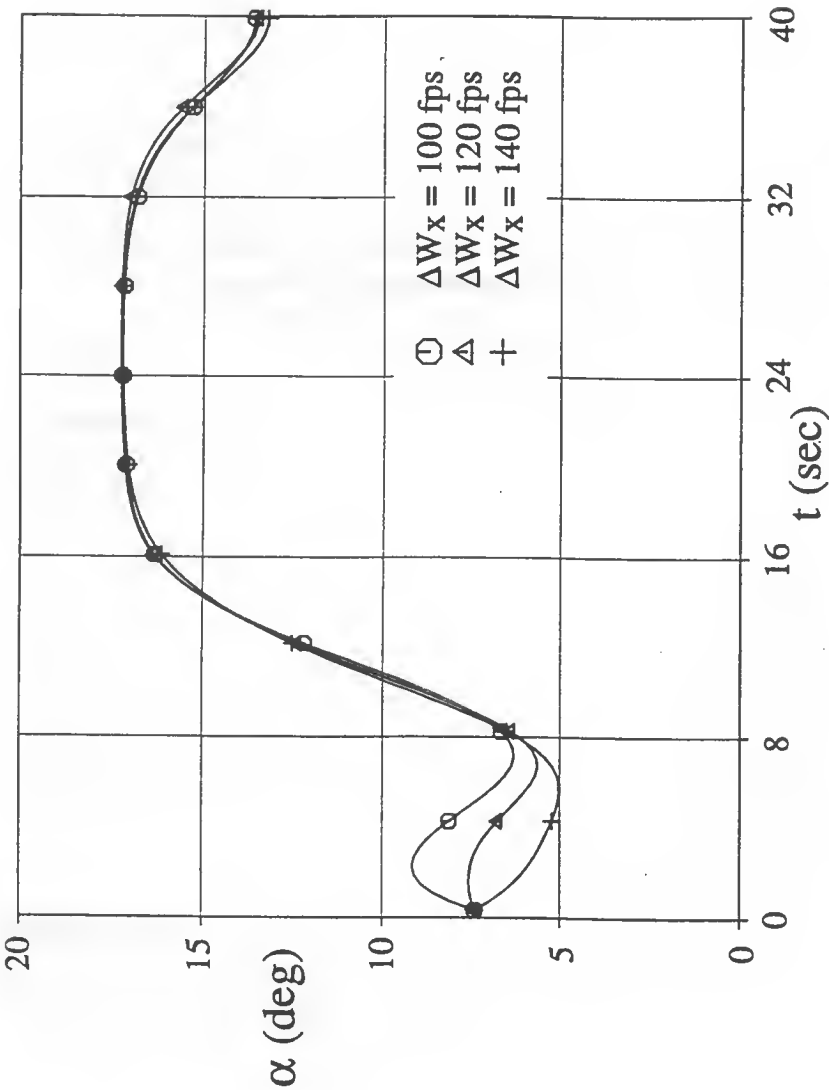


Fig. 8C - Optimal abort landing trajectories, angle of attack vs time.

9. Abort Landing Guidance

Based on the idea of preserving the key properties of the optimal trajectories, while employing local/prior information on the wind acceleration and the downdraft, two guidance schemes were developed at Rice University: (AG) acceleration guidance, based on the relative acceleration; (GG) gamma guidance, based on the absolute path inclination.

9.1. Acceleration Guidance. Let (R1) denote the portion of the unfavorable shear region where descending flight occurs; let (R2) denote the portion of the unfavorable shear region where level flight/mildly ascending flight occurs; let (R3) denote the aftershear region. In feedback control form, the acceleration guidance (AG) is represented by

$$(23a) \quad \alpha - \tilde{\alpha}(V) = K_1(\dot{V}/g + C_1 F), \quad \text{Region (R1),}$$

$$(23b) \quad \alpha - \tilde{\alpha}(V) = K_2(\dot{V}/g + C_2 F), \quad \text{Region (R2),}$$

$$(23c) \quad \alpha - \tilde{\alpha}(V) = K_3(V - C_3 V_0), \quad \text{Region (R3).}$$

In Eqs. (23), K_1, K_2, K_3 are the gain coefficients (positive); C_1, C_2, C_3 are positive constants; α is the guidance angle of attack; $\tilde{\alpha}(V)$ is the nominal angle of attack associated with static equilibrium in the direction normal to the flight path in the absence of shear and downdraft; V is the relative velocity; V_0 is the initial relative velocity; and F is the shear/downdraft factor (12). The meaning of Eqs. (23) is the following: in Region (R1), with $C_1=0$, the relative velocity tends to be constant; in Region (R2), the deceleration increases as the shear/downdraft factor increases; in Region (R3), the relative velocity tends to some fraction of the initial value.

9.2. Gamma Guidance. In feedback control form, the gamma guidance (GG) is represented by

$$(24a) \quad \alpha - \tilde{\alpha}(V) = K_1[\gamma_e - \gamma_{e0}(C_1 h_0/h_* + C_2 F - C_3)], \quad \text{Region (R1),}$$

$$(24b) \quad \alpha - \tilde{\alpha}(V) = K_2[\gamma_e - C_4 \gamma_{er}(1 - C_5 F)], \quad \text{Region (R2),}$$

$$(24c) \quad \alpha - \tilde{\alpha}(V) = K_3(\gamma_e - C_6 \gamma_{er}), \quad \text{Region (R3).}$$

In Eqs. (24), K_1, K_2, K_3 are the gain coefficients (positive); $C_1, C_2, C_3,$

C_4, C_5, C_6 are positive constants; α is the guidance angle of attack; $\bar{\alpha}(V)$ is the nominal angle of attack; γ_e is the absolute path inclination; γ_{e0} is the initial absolute path inclination; γ_{et} is the final absolute path inclination; and F is the shear/downdraft factor (12). The meaning of Eqs. (24) is the following: in Region (R1), the absolute path inclination becomes more negative as the initial altitude increases and the shear/downdraft factor increases; in Region (R2), the absolute path inclination gradually changes from negative values to slightly positive values; in Region (R3), the absolute path inclination tends to some fraction of the final positive value.

9.3. Remark. In both the acceleration guidance and the gamma guidance, suitable switch conditions govern the transition from Region (R1) to Region (R2) to Region (R3). For details, see Ref. 15. Also, for engineering limitations to (23) and (24), see Ref. 15.

9.4. Trajectory Comparison. Using the above schemes, guidance trajectories were computed for the Boeing B-727 aircraft. It was found that both the acceleration guidance (AG) and the gamma guidance (GG) trajectories are close to the optimal trajectories (OT). In addition, the resulting near-optimal trajectories are superior to those arising from maximum angle of attack guidance (MAAG) and constant pitch guidance (CPG). See Fig. 9, which refers to $\lambda = 1.2$, hence $\Delta W_x = 120$ fps. While the MAAG trajectory crashes, the CPG trajectory survives, albeit with minimum altitude about half that of the OT. On the other hand, the AG trajectory is close to the OT, which has good altitude margin. An analogous remark holds for the GG trajectory, which is not shown, since it is nearly identical to the AG trajectory.

9.5. Survival Capability. The merit of a guidance scheme can be assessed using the concept of survival capability. Consider the one-parameter family of wind models (3)-(4). As λ increases, more intense shear/downdraft combinations are generated until a critical value λ_c (hence, a critical value ΔW_{xc}) is found, such that the minimum altitude equals the ground level.

The results are shown in Table 6, which supplies the survival capability ΔW_{xc} for the optimal trajectory and various guidance trajectories. Table 6 also shows the windshear efficiency ratio WER, defined to be the ratio of the survival capability of a particular trajectory to that of the optimal trajectory. Clearly, the windshear efficiency of the OT is 100%, that of the AG trajectory is 96%, that of the GG trajectory

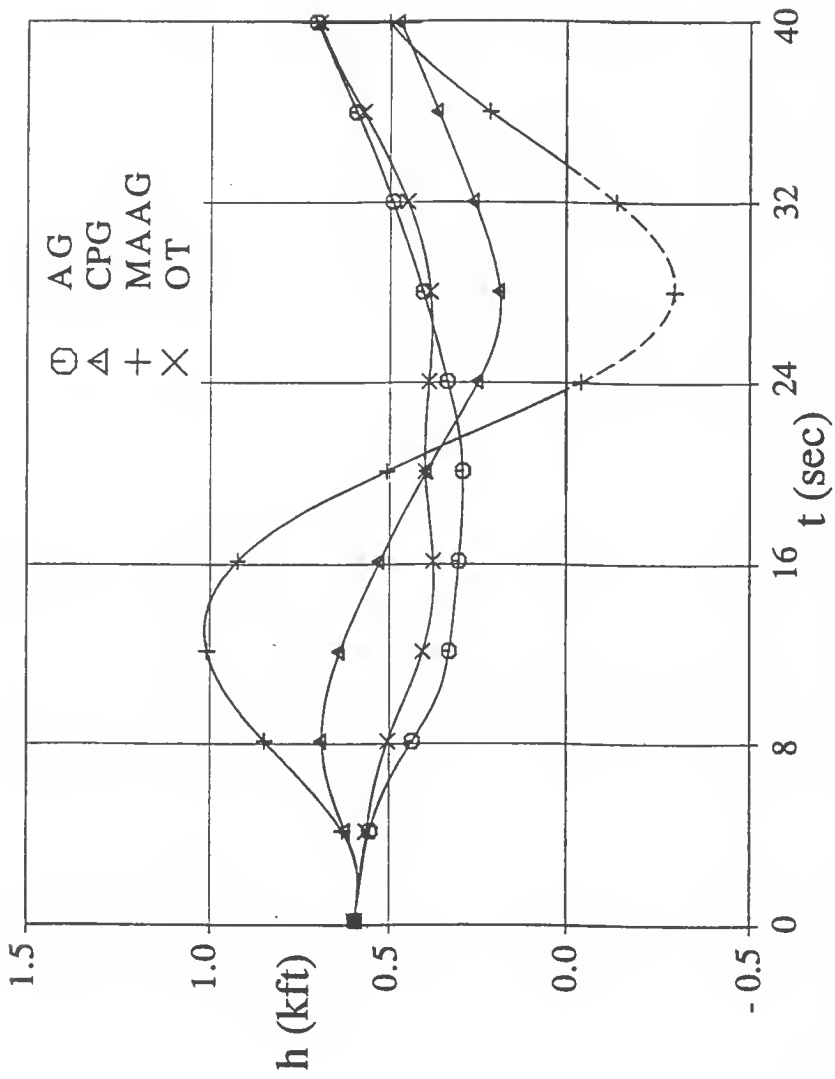


Fig. 9 - Comparison of abort landing trajectories, altitude vs time, $\Delta W_x = 120$ fps.

is 98%, that of the CPG trajectory is 75%, and that of the MAAG trajectory is 44%.

Table 6. Survival capability, abort landing, B-727.

Trajectory	h_0 (ft)	ΔW_{xc} (fps)	WER
OT	600	187.1	1.000
AG	600	179.1	0.957
GG	600	184.2	0.985
CPG	600	139.4	0.745
MAAG	600	81.7	0.437

Transition distance = 4000 ft

10. Penetration Landing Optimization

An alternative to abort landing is the penetration landing maneuver. If the initial altitude of the windshear encounter is relatively high, abort landing is a safer procedure than penetration landing; on the other hand, if the initial altitude is relatively low, the opposite might be true. Indeed, in low-altitude penetration landing, the aircraft might have to traverse only a part of the shear region; however, in low-altitude abort landing, the aircraft might have to traverse the whole of the shear region.

While the abort landing maneuver yields a one-control problem (the power setting is saturated), the penetration landing maneuver yields a two-control problem. The pilot needs both the angle of attack $\alpha(t)$, subject to Ineqs. (13), and the power setting $\beta(t)$, subject to Ineqs. (14), in order to satisfy multiple conditions at touchdown, for instance,

$$(25a) \quad -1000 \leq x_\tau - \tilde{x}_\tau \leq +1000 \text{ ft} ,$$

$$(25a) \quad -50 \leq V_\tau - \tilde{V}_\tau \leq +50 \text{ fps} ,$$

$$(25c) \quad \gamma_{e\tau} = -0.5 \text{ deg.}$$

Here, the tilde denotes nominal conditions.

With the above understanding, optimal penetration landing trajectories

can be determined by minimizing the least-square deviation of the flight trajectory from the nominal trajectory. Thus, the performance index has the form

$$(26a) \quad I_* = \min_{\alpha, \beta} I ,$$

$$(26b) \quad I = \int_0^\tau [h - \tilde{h}(x)]^2 dt .$$

Here, t is the running time; τ is the final time; α is the angle of attack; β is the power setting; x is the horizontal distance; h is the altitude above ground; and $\tilde{h}(x)$ is the nominal altitude. In turn, the function $\tilde{h}(x)$ is computed by assuming that the nominal trajectory includes two parts: the approach part ($h \geq 50$ ft), in which the slope γ_e is constant, $\gamma_e = -3.0$ deg; and the flare part ($h \leq 50$ ft), in which the slope γ_e is a linear function of the horizontal distance.

The Bolza problem represented by the performance index (26), the differential and inequality constraints of Section 5, and suitable boundary conditions is amenable to computational solution by means of the sequential gradient-restoration algorithm; see Ref. 20 for bibliography and recent versions of this algorithm.

10.1. Data and Results. The aircraft under consideration is the Boeing B-727 with landing weight $W = 150000$ lb, gear down, flap deflection $\delta_F = 30.0$ deg, and stick-shaker angle of attack $\alpha_* = 17.2$ deg. At the initial point, the state variables are given (in particular, $h_0 = 600$ ft and $\gamma_{e0} = -3.0$ deg). At the final point, the conditions (25) must be satisfied with $\tilde{x}_\tau = 12130$ ft and $V_\tau = V_0 = 240$ fps, hence

$$(27a) \quad 11130 \leq x_\tau \leq 13130 \text{ ft} ,$$

$$(27b) \quad 190 \leq V_\tau \leq 290 \text{ fps} ,$$

$$(27c) \quad \gamma_{e\tau} = -0.5 \text{ deg} .$$

Optimal trajectories were computed for the above aircraft and the wind model of Section 4. Three values of the wind intensity parameter were employed: $\lambda = 1.0, 1.2, 1.4$. These values correspond to the following wind velocity differences: $\Delta W_x = 100, 120, 140$ fps.

Figure 10 shows the altitude profile $h(x)$, relative velocity profile $V(x)$,

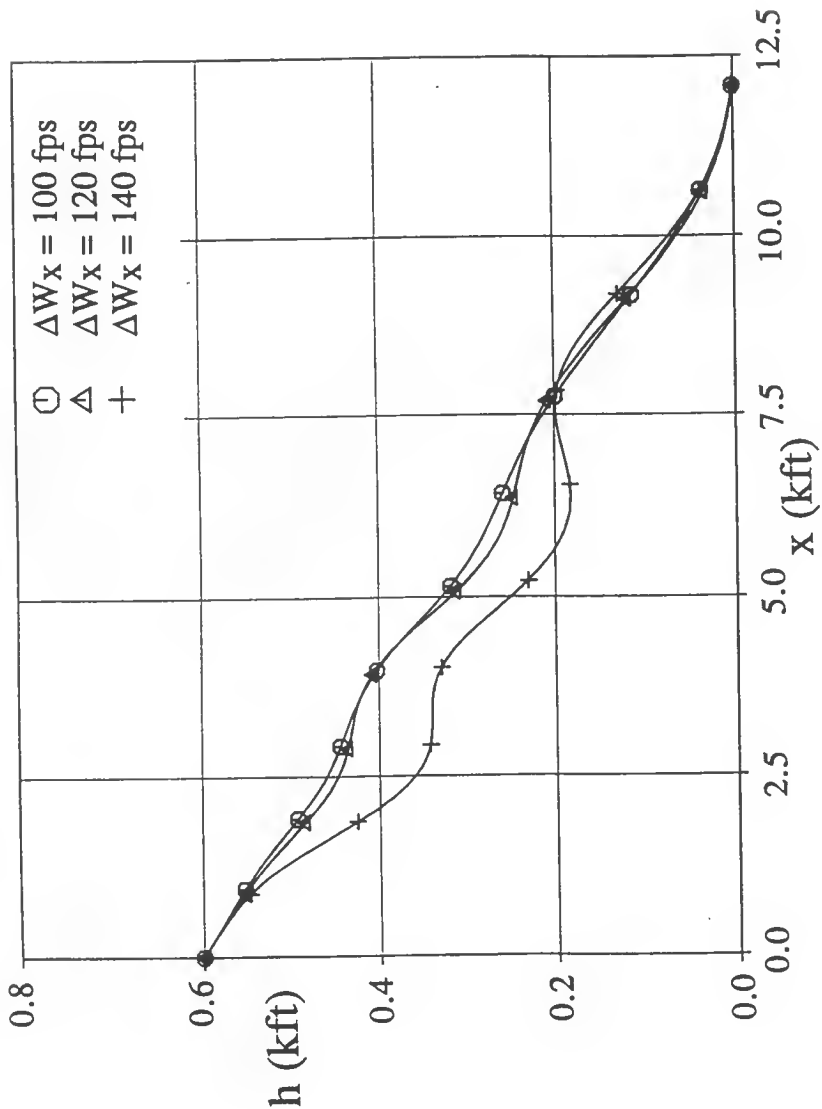


Fig. 10A - Optimal penetration landing trajectories, altitude vs distance.

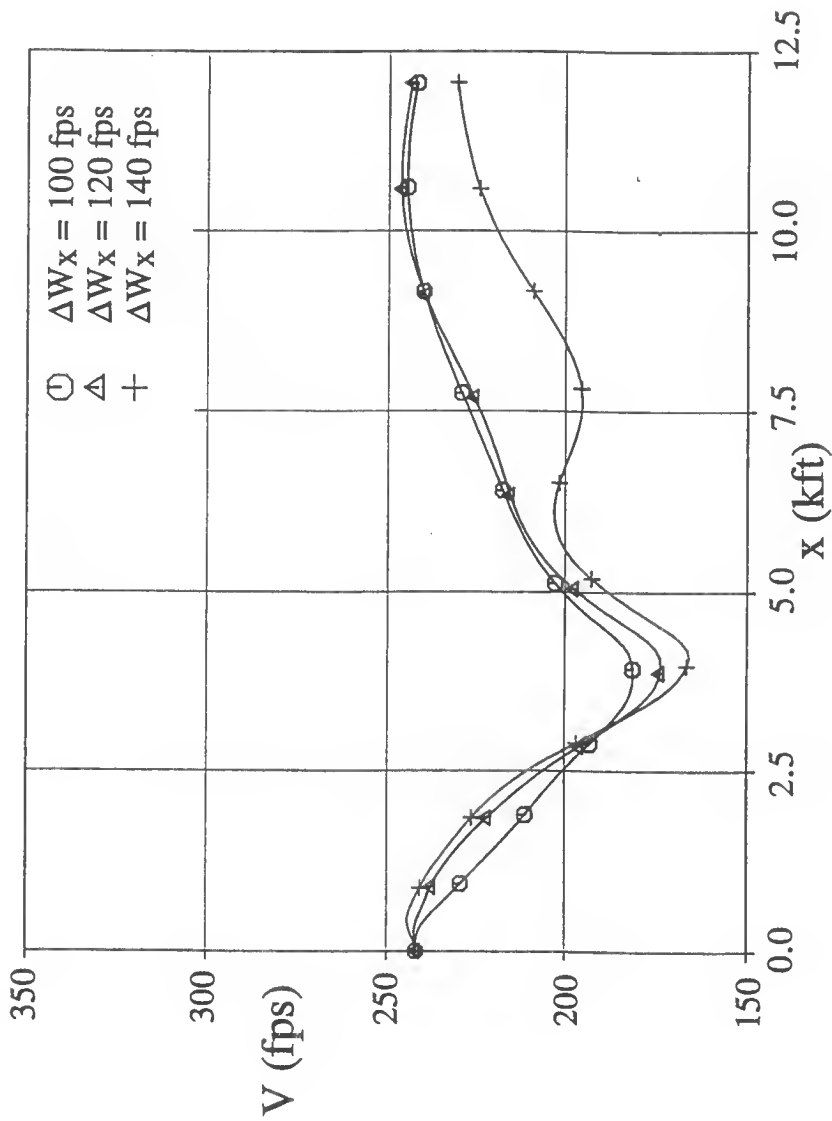


Fig. 10B - Optimal penetration landing trajectories, relative velocity vs distance.

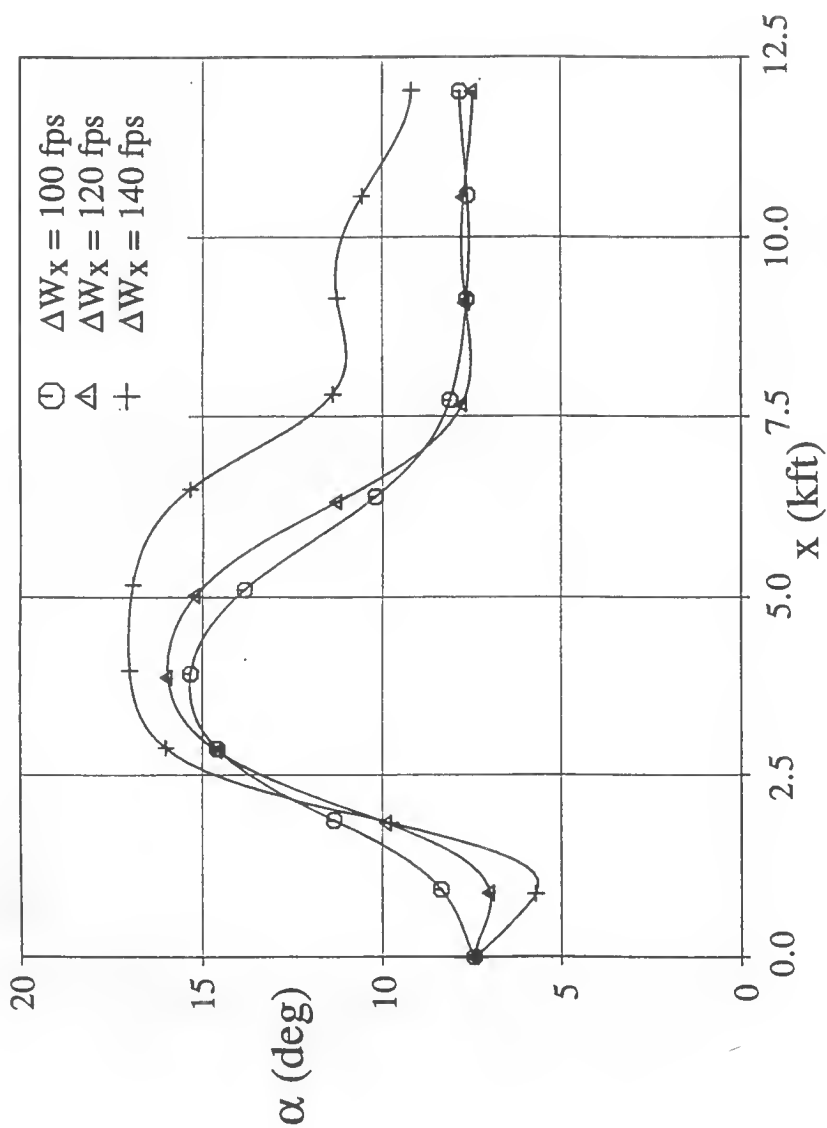


Fig. 10C - Optimal penetration landing trajectories, angle of attack vs distance.

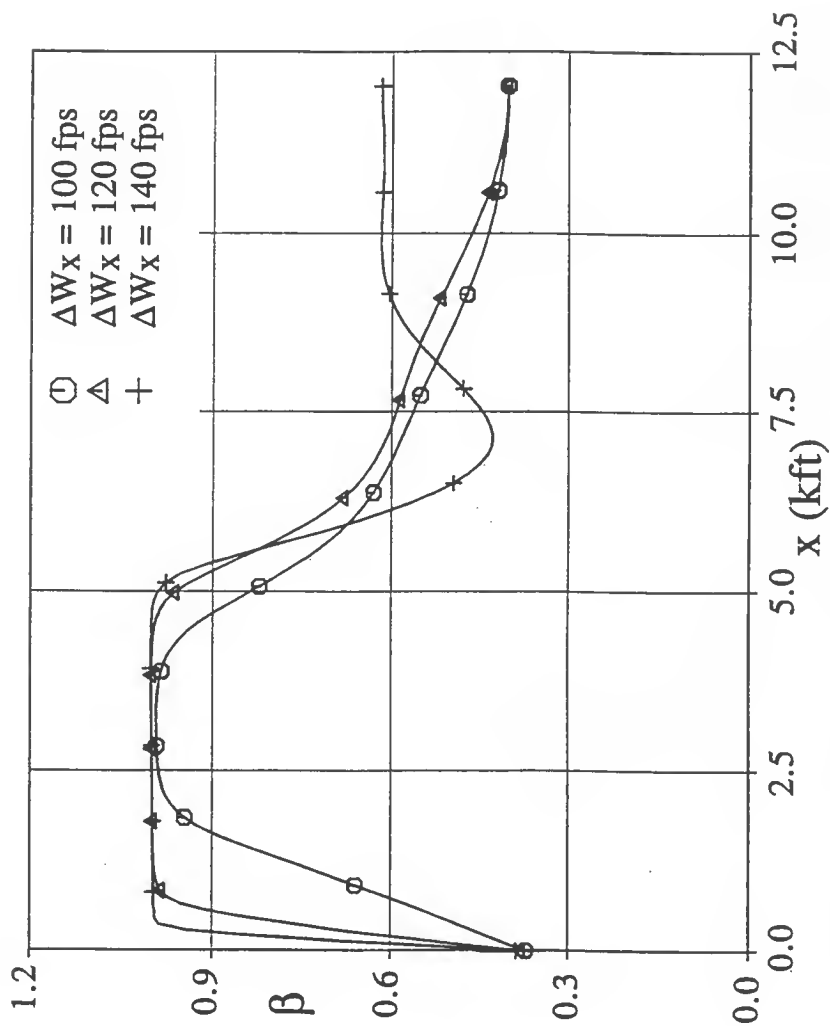


Fig. 10D - Optimal penetration landing trajectories, power setting vs distance.

angle of attack profile $\alpha(x)$, and power setting profile $\beta(x)$. We see that, as ΔW_x increases, the deviations of the optimal trajectory (OT) from the nominal trajectory (NT) increase and can be considerable in the unfavorable shear region; however, these deviations become small in the aftershear region, and the optimal trajectory recovers the nominal trajectory.

10.2. Extensions. The analyses done for $h_0 = 600$ ft and $\dot{\beta}_* = 0.2/\text{sec}$ were extended to several combinations of initial altitude and power setting rate, leading to the following conclusions [Ref. 11]:

- (i) the angle of attack has an initial decrease, which is followed by a gradual, sustained increase; the stick-shaker angle of attack is attained at the end of the shear; in the aftershear region, the angle of attack decreases gradually;
- (ii) initially, the power setting increases rapidly until maximum power setting is reached; this power setting level is maintained throughout the shear region; in the aftershear region, the power setting decreases gradually;
- (iii) the relative velocity decreases in the shear region and increases in the aftershear region; minimum velocity occurs at the end of the shear;
- (iv) depending on the windshear intensity and the initial altitude, the deviations of the flight trajectory from the nominal trajectory can be considerable in the shear region; however, these deviations become small in the aftershear region, and the optimal trajectory recovers the nominal trajectory.

11. Penetration Landing Guidance

A penetration landing guidance (PLG) scheme was developed at Rice University, based on the following ideas: to rely on local/prior information on the wind acceleration, the downdraft, and certain components of the state of the aircraft; to preserve the key properties of the optimal trajectories; and to treat the power setting and the angle of attack as decoupled controls. Because of the separation result established in Refs. 11 and 16, the power setting determination can be based on the velocity, while the angle of attack determination can be based on the absolute path inclination and the glide slope.

11.1. Power Setting. The guidance law must be such that maximum power setting ($\beta = 1$) is maintained in the unfavorable shear region and velocity recovery ($V = V_0$) is ensured in the aftershear region.

In feedback control form, the penetration landing guidance (PLG) is represented by

$$(28a) \quad \beta - \beta_0 = -K_1(V - V_0) + K_2F,$$

$$(28b) \quad K_1 = (1 - \beta_0)/(V_0 - V_l),$$

$$(28c) \quad K_2 = (1 - \beta_0)/F_c.$$

Here, K_1 and K_2 are the gain coefficients (positive); β is the guidance power setting; β_0 is the initial power setting; V is the relative velocity; V_0 is the initial relative velocity; V_l is a lower bound for the relative velocity; F is the shear/downdraft factor (12); and $F_c = 0.125$ is a critical value for the shear/downdraft factor. For details and engineering limitations to (28), see Ref. 16.

11.2. Angle of Attack. The objective of the guidance law is to maintain the flight trajectory close to the nominal trajectory via two signals: the absolute path inclination γ_e and the glide slope γ_g . For the approach part ($h \geq 50$ ft), the glide slope signals are dominant; this is essential to achieving the required touchdown distance. For the flare part ($h \leq 50$ ft), the absolute path inclination signals are dominant; this is essential to achieving the required touchdown absolute path inclination.

In feedback control form, the penetration landing guidance (PLG) is represented by

$$(29) \quad \alpha - \tilde{\alpha}(V) = -K_3 [\gamma_e - \tilde{\gamma}_e(h, F)] - K_4 [\gamma_g - \tilde{\gamma}_g(h, F)].$$

Here, K_3 and K_4 are the gain coefficients (positive); α is the guidance angle of attack; $\tilde{\alpha}(V)$ is the nominal angle of attack; γ_e is the absolute path inclination; $\tilde{\gamma}_e(h, F)$ is the nominal absolute path inclination; γ_g is the glide slope; $\tilde{\gamma}_g(h, F)$ is the nominal glide slope; h is the altitude; and F is the shear/downdraft factor (12). For details and engineering limitations to (29), see Ref. 16.

11.3. Trajectory Comparison. Using the above scheme, guidance trajectories were computed for the Boeing B-727 aircraft. It was found that the penetration landing guidance (PLG) trajectories are relatively close to the optimal trajectories (OT). In addition, the resulting near-optimal

trajectories are superior to those arising via fixed-control guidance (FCG) and autoland guidance (ALG) in terms of the ability to meet the touch-down requirements and in terms of survival capability in strong-to-severe windshears. See Fig. 11, which refers to $\lambda = 1.2$, hence $\Delta W_x = 120$ fps. Both the FCG trajectory and the ALG trajectory crash; however, the PLG trajectory survives the windshear encounter, just as does the OT.

12. Summary, Conclusions, and Recommendations

Windshears originating from downbursts have been the cause of many aircraft accidents in the last two decades. To offset the windshear threat, both avoidance systems and recovery systems have been developed.

In this paper, the Rice University effort on recovery systems is reviewed. First, optimal trajectories are discussed for three basic flight conditions: take-off, abort landing, and penetration landing. Because the optimal trajectories require global information on the wind field, they cannot be flown in real time. However, they constitute a benchmark that one should strive to approach in flight.

Starting from the optimal trajectories, guidance schemes have been developed which are based on the following idea: to approximate the key properties of the optimal trajectories, while employing local/prior information on the wind acceleration, the downdraft, and certain components of the state of the aircraft. Noteworthy among the guidance schemes are these: for take-off, acceleration guidance and gamma guidance; for abort landing, acceleration guidance and gamma guidance; for penetration landing, penetration landing guidance. The survival capability of the resulting trajectories is close to that of the optimal trajectories and compares favorably with that of alternative trajectories.

While much progress has been accomplished thus far, much remains to be done. Consider for instance the area of detection, which is essential to avoidance. Clearly, the shear/downdraft factor (12) can be employed as a signal to detect the windshear. In a real situation, this signal might be blurred by the presence of free-stream turbulence; in turn, this might result in undesirable nuisance warnings if the magnitude of the shear factor due to turbulence is temporarily larger than that due to windshear. Therefore, proper filtering is necessary prior to the use of the shear/downdraft factor (12) in detection and guidance.

An alternative to the use of (12) is to employ the wind difference signal for detection and guidance. This signal is more stable than the shear/downdraft factor signal with respect to interference effects due

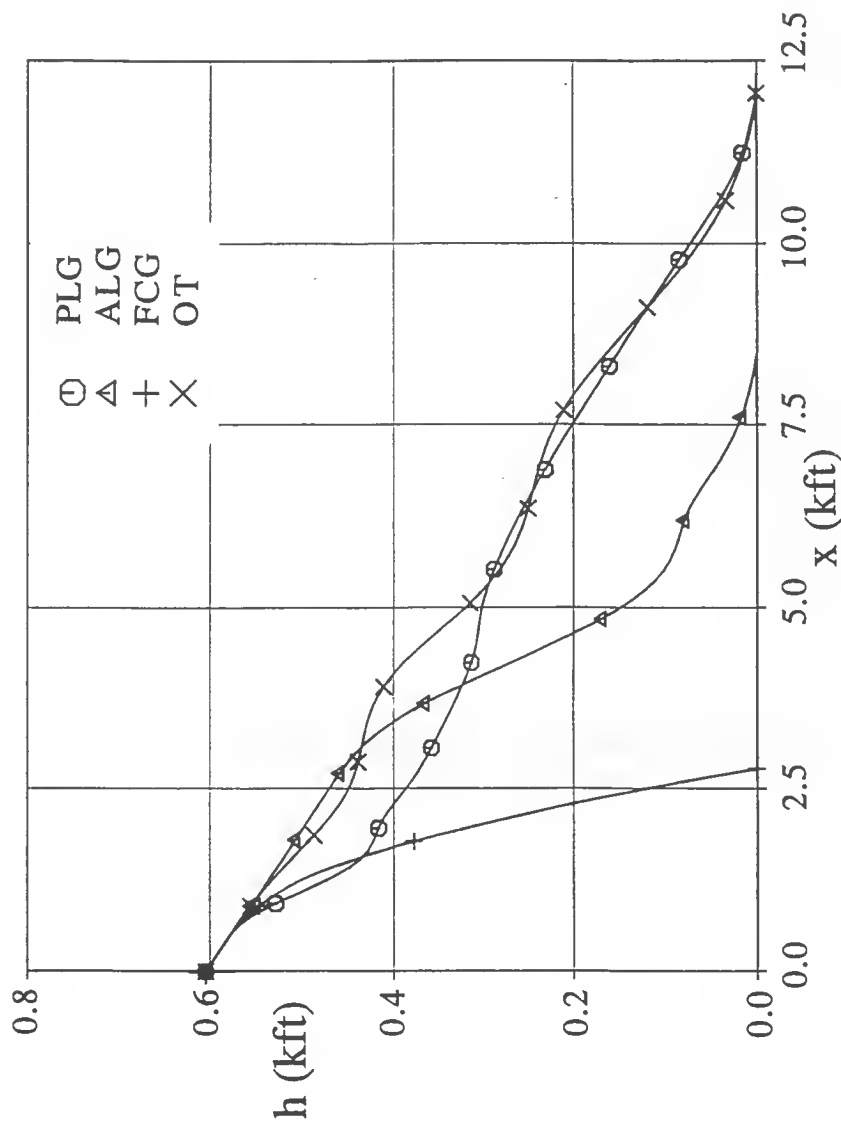


Fig. 11 - Comparison of penetration landing trajectories, altitude vs distance, $\Delta W_x = 120$ fps.

to free-stream turbulence; yet, the wind difference signal can be determined using the same aerodynamic/inertial instrumentation necessary for (12); see the recent work on detection by the Aero-Astronautics Group of Rice University [Refs. 22-24].

A byproduct of the point of view taken in Refs. 22-24 is that it appears possible to combine the warning function and the recovery function into a single system. Thus, the development of an advanced windshear control system, endowed with greater stability vis-a-vis interference effects due to free-stream turbulence, is in sight.

It must be noted that the recovery systems (guidance systems) discussed in this paper are rather conservative: they are specifically developed for the case of an unfavorable wind gradient. It must also be noted that any adverse wind gradient (inner core of a downburst) is both preceded and followed by a favorable wind gradient. The advanced windshear recovery system must not only react in a near-optimal way to adverse wind gradients, but must exploit to the best advantage of the aircraft favorable wind gradients. This means that, in an increasing headwind scenario, kinetic energy must be increased; conversely, in a decreasing tailwind scenario, potential energy must be increased. This requires that not only the current windshear signals be measured, but that previous windshear signals be recorded and memorized, such that favorable wind gradients can be detected and used.

To sum up, it is felt that an advanced windshear recovery system, endowed with the properties described above, should improve considerably the survival capability of an aircraft in severe windshear.

REFERENCES

- [1] ANONYMOUS N.N., *Low Altitude Windshear and Its Hazard to Aviation*, National Academy Press, Washington, DC, 1983.
- [2] ANONYMOUS N.N., *Windshear Training Aid*, Vols. 1 and 2, Federal Aviation Administration, Washington, DC, 1987.
- [3] FUJITA T.T., *The Microburst*, SMRP Research Paper No. 210, Department of Geophysical Sciences, University of Chicago, 1985.

- [4] FUJITA T.T., *DFW Microburst*, SMRP Research Paper No. 217, Department of Geophysical Sciences, University of Chicago, 1986.
- [5] MIELE A., WANG T. and MELVIN W.W., *Optimal Take-Off Trajectories in the Presence of Windshear*, Journal of Optimization Theory and Applications, Vol. 49, No. 1, pp. 1-45, 1986.
- [6] MIELE A., WANG T. and MELVIN W.W., *Guidance Strategies for Near-Optimum Take-Off Performance in a Windshear*, Journal of Optimization Theory and Applications, Vol. 50, No. 1, pp. 1-47, 1986.
- [7] MIELE A., WANG T., MELVIN W.W. and BOWLES R.L., *Maximum Survival Capability of an Aircraft in a Severe Windshear*, Journal of Optimization Theory and Applications, Vol. 53, No. 2, pp. 181-217, 1987.
- [8] MIELE A., WANG T. and MELVIN W.W., *Quasi-Steady Flight to Quasi-Steady Flight Transition in a Windshear: Trajectory Optimization and Guidance*, Journal of Optimization Theory and Applications, Vol. 54, No. 2, pp. 203-240, 1987.
- [9] MIELE A., WANG T., TZENG C.Y. and MELVIN W.W., *Optimal Abort Landing Trajectories in the Presence of Windshear*, Journal of Optimization Theory and Applications, Vol. 55, No. 2, pp. 165-202, 1987.
- [10] MIELE A., WANG T. and MELVIN W.W., *Optimization and Acceleration Guidance of Flight Trajectories in a Windshear*, Journal of Guidance, Control, and Dynamics, Vol. 10, No. 4, pp. 368-377, 1987.
- [11] MIELE A., WANG T., WANG H. and MELVIN W.W., *Optimal Penetration Landing Trajectories in the Presence of Windshear*, Journal of Optimization Theory and Applications, Vol. 57, No. 1, pp. 1-40, 1988.
- [12] MIELE A., WANG T., MELVIN W.W. and BOWLES R.L., *Gamma Guidance Schemes for Flight in a Windshear*, Journal of Guidance, Control, and Dynamics, Vol. 11, No. 4, pp. 320-327, 1988.
- [13] MIELE A., WANG T. and MELVIN W.W., *Quasi-Steady Flight to Quasi-Steady Flight Transition for Abort Landing in a Windshear: Trajectory Optimization and Guidance*, Journal of Optimization Theory and Applications, Vol. 58, No. 2, pp. 165-207, 1988.
- [14] MIELE A., WANG T., TZENG C.Y. and MELVIN W.W., *Abort Landing Guidance Trajectories in the Presence of Windshear*, Journal of the Franklin Institute, Vol. 326, No. 2, pp. 185-220, 1989.

- [15] MIELE A., WANG T., MELVIN W.W. and BOWLES R.L., *Acceleration, Gamma, and Theta Guidance for Abort Landing in a Windshear*, Journal of Guidance, Control, and Dynamics, Vol. 12, No. 6, pp. 815-821, 1989.
- [16] MIELE A., WANG T. and MELVIN W.W., *Penetration Landing Guidance Trajectories in the Presence of Windshear*, Journal of Guidance, Control, and Dynamics, Vol. 12, No. 6, pp. 806-814, 1989.
- [17] ZHU S. and ETKIN B., *Model of the Wind Field in a Downburst*, Journal of Aircraft, Vol. 22, No. 7, pp. 595-601, 1985.
- [18] IVAN M., *A Ring-Vortex Downburst Model for Flight Simulation*, Journal of Aircraft, Vol. 23, No. 3, pp. 232-236, 1986.
- [19] WANG H., *Optimization of Flight Trajectories in a 3D-Model of Windshear Flow Field*, PhD Thesis, Rice University, Houston, Texas, 1992.
- [20] MIELE A., *Optimal Trajectories of Aircraft and Spacecraft*, Aircraft Trajectories: Computation, Prediction, and Control, Edited by A. Benoit, AGARDograph No. 301, AGARD/NATO, Paris, France, Vol. 1, pp. 201-256, 1990.
- [21] MIELE A., WANG T., WANG H. and MELVIN W.W., *Overview of Optimal Trajectories for Flight in a Windshear*, Control and Dynamics Systems, Advances in Theory and Application, Edited by C.T. Leondes, Vol. 34, pp. 81—123, 1990.
- [22] MIELE A., WANG T., WANG H. and MELVIN W.W., *Introduction to an Advanced Detection System for Flight in a Windshear*, Rice University, Aero-Astronautics Report No. 255, 1991.
- [23] MIELE A., WANG T. and MELVIN W.W., *Real-Time Wind Identification Technique*, Rice University, Aero-Astronautics Report No. 256, 1991.
- [24] MIELE A., WANG T. and MELVIN W.W., *Advanced Windshear Detection Technique*, Rice University, Aero-Astronautics Report No. 257, 1991.



La sicurezza nel progetto degli aeromobili e dei veicoli spaziali

Ettore ANTONA

A Giuseppe Gabrielli, per onorare la Sua memoria e per progredire, per quanto mi è dato, nello sviluppo delle Sue idee e del Suo insegnamento.

1. Introduzione

Il presente lavoro propone riflessioni ai livelli teorico e metodologico, nell'intento di contribuire a ulteriori sviluppi del pensiero sulla sicurezza e sui concetti collegati. Esso, quindi, non deve essere considerato fonte di indicazioni, di natura applicativa, per le quali vale l'aggiornata normativa.

Un nuovo interesse a livello speculativo per i problemi della sicurezza nell'attività aerospaziale si sviluppa ormai da più di un ventennio. Esso va considerato come una ideale continuità rispetto al passato, nel quale già si era imposta anche in ambito accademico una cultura, che vede l'ingegneria dal versante della sintesi nella realizzazione, dove è elemento tra i dominanti il quadro dei compiti dei vari operatori, illuminato da un pensiero scientifico sui rischi e sul loro controllo.

A livello filosofico l'idea centrale della attuale concezione della sicurezza è rappresentata dalla accettazione della sua «*natura probabilistica*», nella necessità di fronteggiare i rischi con tutta una serie di cautele, siano esse di natura deterministica, probabilistica o mista. Se da una parte, all'attuale grado di sviluppo della tecnica, risulta pressoché impossibile la valutazione di un dato onnicomprensivo della sicurezza stessa, specie per taluni aspetti dei rischi, la corretta impostazione probabilistica risulta di grande vantaggio nella chiarificazione dei concetti.

La tecnica aeronautica e spaziale ha dato un enorme contributo allo sviluppo di questa concezione. Molti altri campi di attività ne sono oggi interessati o ne dovrebbero essere interessati.

L'approccio alla sicurezza, che caratterizza le attività aeronautiche e spaziali, necessita di approfondimenti e analisi, da svolgersi principalmente sul piano teorico. Lo scopo ultimo è quello di appurare le capacità di ottenere per un qualunque sistema «macchina-impiego-pilota» la valutazione delle probabilità di successo. A maggior ragione è utile tenerla per la sola macchina.

Un altro scopo è quello di stabilire basi teoriche per la valutazione delle probabilità globali di successo e per la critica della attendibilità dei dati disponibili.

Le considerazioni cui si è indotti dalla tecnica aeronautica, messe a confronto con le rimanenti tecniche, fanno concludere che, a maggior ragione, sono per essa necessarie analisi fondate su basi scientifiche, sia per determinare le probabilità di funzionamento dei sistemi, sia per vagliare i dati, su cui si fondano tali valutazioni o per acquisire i dati stessi.

Compito degli studi e delle ricerche deve essere, allora, quello di costruire la capacità di valutare preventivamente la probabilità di funzionamento di un sistema e, tenuto conto della diversa pericolosità dei vari eventi, di valutare preventivamente il rischio connaturato con un sistema.

Se si esamina l'evoluzione della tecnica, la sicurezza, molto perseguita a livello tecnico, quale argomento di ricerca e studio, pur presente, appare piuttosto di impegno inferiore a quello ipotizzabile.

La maggior parte degli sforzi tendenti ad assicurare alle realizzazioni dell'uomo la opportuna «sicurezza», risiedono principalmente nella normativa e in altre attività, più vicine alla pratica che alla teoresi.

Questo non vuol dire che gli uomini di scienza non avvertano abbastanza il problema della sicurezza; significa solamente che, forse istintivamente, a tale argomento si è più inclini ad associare l'ambito tecnico che non quello scientifico.

La distinzione fra ambito tecnico e ambito scientifico, sulla quale spesso abbiamo avuto occasione di essere indotti a riflettere, ci offre ancora una volta la possibilità di vedere come, sotto la spinta di necessità, quale è certamente la sicurezza, la tecnica costruisce conoscenza e cultura anche prima che esse diventino oggetto di studi della scienza, così aprendone nuovi capitoli. Diversamente da quanti vedono in ciò una sorta di primogenitura della tecnica, ritengo occorra tener presente i grandi contributi che anche la scienza ha dato nell'aprire nuove tecnologie.

Anche sull'argomento della sicurezza sono convinto che un impegno scientifico approfondito sia opportuno, anzi necessario, per la penetrazione culturale che ne potrà derivare, oltretutto, ovviamente, per i contributi specifici.

La definizione di «sicurezza» è: «L'assenza di quelle condizioni che possono causare morte, ferimenti, malattie professionali, il danneggiamento oppure la perdita di equipaggiamenti o di beni» (MIL STD 882 B cui rimanda la MIL 721 C). Questa definizione ha recentemente sostituito la precedente: «La conservazione della vita umana e della integrità

delle persone, nonché la prevenzione dei danni alle macchine, nel rispetto dei requisiti della missione» (MIL 721 B)⁽¹⁾.

La «sicurezza» così definita è un problema della ingegneria, senza riferimenti alla sua natura, sia essa deterministica o probabilistica.

La sicurezza ha avuto aspetti probabilistici sin dai primordi dell'aeronautica, in particolare per quanto riguarda le caratteristiche dei materiali e le valutazioni dei carichi di volo.

Durante l'evoluzione dell'attività aeronautica e spaziale, gli aspetti probabilistici hanno accresciuto fortemente la loro importanza, fino ad essere coinvolti anche nelle analisi (statistiche) delle probabilità di rottura dei vari componenti e, più recentemente, dei carichi di rottura delle strutture, delle propagazioni delle cricche, delle capacità dei metodi di controllo non distruttivi, e così via.

Il livello complessivo di «sicurezza» è il risultato di un grande numero di aspetti, molti dei quali, da un punto di vista teorico, possono essere analizzati solamente su base statistica.

Perciò possiamo dire in termini accademici che la sicurezza è la probabilità di non avere «rischi» (cioè eventi tali da potenzialmente procurare danni) durante un prefissato impiego. Ne consegue che, da un punto di vista generale, la determinazione del livello di sicurezza ottenibile può essere fatta solo su base statistica.

D'altra parte, nelle pratiche di progetto, la sicurezza è perseguita per mezzo di prescrizioni, che sono in definitiva di natura probabilistica, in quanto criteri deterministici si applicano (molto spesso) a dati e valori di origine probabilistica.

In effetti, la sicurezza complessiva ottenibile (o ottenuta) è di calcolo estremamente difficile, allo stato attuale dello sviluppo delle conoscenze e dei dati, in specie se devono essere tenuti in conto gli eventi meteorologici e naturali e gli errori umani o per quei componenti per cui è estremamente prossima all'unità.

Per gli scopi dei futuri studi teorici e sperimentali accanto all'adeguamento alla evoluzione tecnica, ci dovrà essere l'adeguatezza delle formulazioni teoriche e dei dati pratici, con il proposito di valutare la sicurezza sotto forma di una stima globale.

L'esistenza di un livello complessivo di sicurezza, risultante da tutte le prescrizioni di sicurezza, considerate in un tutt'uno, anche se non chia-

⁽¹⁾ Le così dette «Norme Mil», indicate con una abbreviazione di «Military» sono emesse dal Dipartimento della Difesa, ma sono ampiamente usate anche in applicazioni civili, rappresentando un contributo essenziale alla organizzazione delle tecnologia. Spesso sono citate anche in capitoli e specifiche.

ramente valutabile allo stato presente delle conoscenze, e la accettazione di tale livello, quale obbiettivo dell'intero complesso di prescrizioni, offre una visione chiara del problema, nel presente momento della attività aeronautica e spaziale.

I dati desunti dall'esercizio dei vari mezzi di trasporto, rappresentano una fonte «certa» di informazioni, dalle quali iniziare l'analisi delle probabilità di incidenti. Il dato più certo è quello che prescinde dalle cause, ovviamente. Via via che si desidera analizzare in funzione delle cause gli incidenti avvenuti, si hanno via via più incertezze, perché ciò comporta l'attribuzione di ogni incidente ad una causa determinata.

La sicurezza si colloca sul piano probabilistico; è quindi immediato comprendere come, a livello di strumenti di analisi e di progetto, essa abbia forti legami con l'affidabilità e i concetti collegati, quali la disponibilità, la manutenibilità e altri. È per queste ragioni che nel presente lavoro, nel presentare un panorama di concetti teorici e dei relativi metodi di approccio di un campo probabilistico, tutti questi concetti sono considerati in una unica sezione.

2. Strategie per la sicurezza

La «sicurezza» dell'intero sistema, quindi, è un concetto che può essere scientificamente analizzato solamente sul piano probabilistico, ed è giocoforza accettare che una probabilità di azzardo⁽²⁾ sia ineliminabile. Ciò è vero per la sicurezza tenuto conto dell'intero novero degli elementi contribuenti (fattore umano, condizioni ambientali, guasti o rotture di ogni apparato o sistema), ma a ben guardare si applica anche a ogni singolo contributo alla causa di rischio⁽²⁾.

Il rischio accettabile nella visione globale della sicurezza è un concetto che allo stato della tecnica risulta astratto, perché di non affidabile valutazione il rischio connotato con un determinato velivolo (o una determinata flotta). Le Norme, con le loro prescrizioni e raccomandazioni, innumerevoli e dettagliate, tendenti a prendere precauzioni per una enorme quantità di situazioni ed evenienze, costituiscono, a mio avviso, l'unica fonte di definizione, sia pure in modo indiretto, della «accettabile

⁽²⁾ Con «azzardo» si intende qui «situazione pericolosa», cui non è detto segnano danni, cioè quel che in inglese significa «hazard». L'inglese «risk» viene invece usato per indicare le probabilità di un danno in combinazione con l'entità del danno stesso. Questo concetto può essere espresso con «rischio».

probabilità di azzardo». La normativa è in continua evoluzione e tiene conto di tutti gli eventi dell'attività aeronautica nonché di tutti i risultati ottenuti dagli studiosi per via teorica o sperimentale.

La strategia applicata costantemente dalle Norme è quella di individuare, per le innumerevoli condizioni di cui ci si deve preoccupare, «sicuri» (cioè di una certa improbabilità nel generare azzardi) limiti di impiego e poi di evitare (entro limiti probabilistici) che tali limiti siano superati. Sia la probabilità di generazioni di azzardi, sia quella di superare i limiti predetti, che vengono usualmente prese in considerazione, tengono conto delle ulteriori possibilità di intervento, specie del pilota, per ovviare alle conseguenze.

L'attività aeronautica dal punto di vista della sicurezza è oggetto di attenzioni a livello scientifico, tecnico e operativo, le quali, in conseguenza della estensione mondiale del teatro operativo e dell'importanza delle conseguenze degli incidenti, hanno di fatto e per molti versi anche formalmente base mondiale. Si può dire in sintesi che i provvedimenti che si introducono per ottenere sicurezza, le osservazioni sperimentali, le analisi teoriche dei problemi e le normative che li impongono sono gli stessi in tutti i paesi del mondo. Ciò non basta però a sintetizzare lo stato di collaborazione e di sintonia che sussiste fra i vari paesi, costituenti di fatto un ambito nel quale qualunque esperienza va a vantaggio di tutti e chiunque può sollevare problemi o aprire processi di critica e approfondimento dei criteri e dei dati a fondamento della sicurezza.

Esistono in ciascun paese di importanza aeronautica Organismi pubblici cui è demandato il compito di elaborare Regolamenti sulla sicurezza e di verificarne l'applicazione. In sintesi il problema della sicurezza è sovrastato dalle seguenti verità ineludibili:

- 1) non è possibile ottenere la certezza di un funzionamento esente in assoluto da rischi;
- 2) la valutazione del rischio insito in una attività può essere fatta solamente in termini probabilistici;
- 3) se la probabilità di azzardo viene ridotta a essere estremamente remota, da un punto di vista pratico si può considerare di aver ottenuta la certezza.

Ciò comporta la necessità di analizzare le probabilità del manifestarsi di eventi semplici o complessi.

Il corretto e soddisfacente svolgimento di una missione dipende dal corretto funzionamento di apparati e dalla corretta manovra, tenuto conto delle condizioni in cui il volo si svolge.

Individuate le combinazioni di guasti, errori e condizioni che indipendentemente fra loro possono essere causa di azzardi, la probabilità

totale di rischio è la somma delle probabilità delle singole combinazioni.

La probabilità del verificarsi di una combinazione può essere di determinazione assai elaborata. Riferendoci a titolo di esempio al caso di eventi che provocano azzardi solamente se si verificano concomitantemente, tale probabilità può essere calcolata facendo ricorso al concetto complementare di probabilità di *non verificarsi* della combinazione. Questa risulta il prodotto delle corrispondenti probabilità di *non verificarsi* dei singoli eventi.

Alcune combinazioni si identificano semplicemente con un singolo guasto o rottura: ad esempio una esplosione che pregiudichi i comandi. Altre risultano da concomitanze di guasti ed errori, sì che è possibile ottenere che il rischio connesso sia estremamente remoto, attraverso la impostazione di limiti opportuni, alla probabilità di manifestarsi dei vari eventi della combinazione.

L'attività aeronautica, nel campo del trasporto civile, è supportata da una serie di accordi e intese internazionali, che consentono i voli su scala mondiale e i necessari interscambi.

Le strategie per la sicurezza hanno quindi ormai una base comune su scala mondiale e si può dire che tale base si avvalga di esperienze, sforzi e pratiche di tutti i paesi, ottenendo livelli di sicurezza altrimenti non raggiungibili, specie dei paesi via via meno avanzati.

Una valutazione della sicurezza sotto forma di valori numerici è possibile solo in campo probabilistico.

Ad esempio i dati a posteriori sulla sicurezza dei mezzi di trasporto sono tutti di natura statistica. Ciò rimane vero anche se ci si limita alle cause di rischio connesse con la macchina in sé e per sé.

Di fatto ogni componente è oggetto di una probabilità di guasto in un dato tempo di funzionamento. Di fronte a questa realtà le strategie per la sicurezza si possono articolare in vari modi fra i due estremi:

- qualitativo, consistente nel duplicare, triplicare, ecc. ogni componente;
- quantitativo, consistente nell'abbassare la probabilità di guasto fino al necessario.

Non interessa qui una disquisizione sui vantaggi e svantaggi dei due modi. Interessa che in ogni caso il risultato globale è esprimibile solamente in termini probabilistici e non otterrà, in termini rigorosi, le «certezze» nella sicurezza.

Della natura probabilistica della sicurezza sotto la guida della dottrina, in aeronautica ci si è resi conto gradualmente ma irreversibilmente.

Anche in altre tecniche, quali ad esempio principalmente gli impianti nucleari e le costruzioni civili, si è sempre posto o si va ormai ponendo il problema della sicurezza in termini probabilistici.

È impossibile, riferendoci in particolare all'aeronautica, ottenere un

valore unitario della probabilità di non avere azzardi, cioè di avere sicurezza totale.

Ciò è vero anche per i rischi dipendenti dalla macchina ed è dovuto alla aleatorietà sia delle situazioni e condizioni in cui la macchina può venirsi a trovare sia delle capacità intrinseche dei materiali o dei componenti di sopportare tali situazioni e condizioni.

Così come non è possibile ottenere un pilota che non commetta errori in assoluto, è impossibile concepire e realizzare una macchina senza rischi. L'unica via è rendere errori e azzardi estremamente improbabili.

In particolare, per quanto riguarda le condizioni nelle quali il volo potrebbe svolgersi, è impossibile determinare limiti di massima onerosità, data la natura aleatoria delle condizioni stesse. Ad esempio: i moti dell'aria, che combinati con la velocità di volo, creano carichi sulle strutture, sono stati indagati in natura con sperimentazioni in volo. Ma per quanto estese le sperimentazioni siano state, esse rappresentano una campionatura della realtà e non possono certamente escludere che si possano verificare in aria situazioni più severe delle più severe incontrate nella sperimentazione.

Anche le analisi teoriche conducono a modelli nei quali non v'è limite alla severità delle situazioni originatesi in aria: unica risultanza utilizzabile è che le situazioni più sono severe e meno sono probabili, seguendo leggi che sperimentazioni e teoria contribuiscono a determinare. Unica possibile scelta è quindi la individuazione di una probabilità accettabile, con conseguente determinazione del corrispondente livello di severità. Quest'ultimo si considera verificabile in volo e si impone quale criterio per la costruzione che la struttura possa sopportare, senza cedere, un livello convenientemente maggiore. Vi sono però situazioni meteorologiche, nelle quali è relativamente elevata la probabilità di incontrare condizioni più severe di quelle prese in considerazione della Normativa con il processo logico predetto. Tali situazioni meteorologiche devono essere evitate e la strategia per la sicurezza si compone in un tutt'uno, della sopportabilità delle condizioni previste dalla Normativa per la costruzione e della pianificazione e condotta dei voli, che eviti delle condizioni o, «estrema ratio» le loro conseguenze. Nella «condotta dei voli» vanno incluse anche tutte le possibili azioni, quali adozione di velocità opportune, che riducano gli oneri, cambiamenti di quote e di rotta, che consentano di padroneggiare la situazione nel caso ci si trovi in condizioni che avrebbero dovuto essere evitate con altri provvedimenti. Lo scopo è quello di ottenere una probabilità di azzardo complessiva estremamente remota, paragonata per esempio ai rischi della vita ordinaria, e che comunque ai fini pratici possa essere accettata come pratica «certezza» di sicurezza.

Nel termine «condotta di pilotaggio» si include anche tutto quel che

attiene alle misure che si possono prendere «estrema ratio» con il modificare le velocità, le quote, le rotte nel caso una situazione severa debba esser fronteggiata, e che insieme a tutti gli altri accorgimenti entra a far parte integrante della fronteggiabilità del rischio.

Fanno, dunque, parte integrante di questo complesso dispositivo di sicurezza, l'unico attuabile, vari fattori.

Il controllo del traffico comprende (fra l'altro) il controllo meteo, che è in grado di individuare e comunicare con sufficiente anticipo le condizioni pericolose per la navigazione: gran parte delle condizioni eccedenti i limiti della certificazione di aeronavigabilità sono cioè prevedibili e prevenibili.

Talune condizioni meteorologiche sono visibili ai radar di terra e di bordo, e possono essere evitate. Quando il pilota individua condizioni di pericolo, o ne riceve segnalazione, deve in qualunque momento decidere di mutare o invertire la rotta.

Le condizioni di certificazione contemplate nei Regolamenti e a cui si fa riferimento nella costruzione dell'aeromobile, si inseriscono in questo complesso e sono il risultato di prove e ricerche svolte da decenni su scala mondiale.

Quell'imponderabile che possa ancora sfuggire alla sequenza di barriere poste all'azzardo deve intendersi come rischio connesso con lo svolgimento dell'attività.

Nella definizione MIL della sicurezza, non si parla di probabilità perché la sicurezza va perseguita in ogni modo, anche se le precauzioni non vengono prese apparentemente in termini probabilistici o solo probabilistici.

In ogni caso, tali precauzioni producono effetti che possono essere valutati solo a livello probabilistico, da un punto di vista quantitativo.

Le precauzioni da assumersi nella progettazione e nella realizzazione, nella manutenzione, nell'esercizio e in ogni altra fase o aspetto dell'attività, in altre parole i livelli di sicurezza da realizzare, sono fissati nei Regolamenti e Normative, che, nelle forme che ciascun paese ha voluto darsi, esprimono la volontà della collettività.

È giusto che i tecnici e la dottrina pongano il problema nei termini comprensibili, ma che sia la collettività a fissare i livelli di sicurezza.

Regolamenti e Normative, introducono dati, valori, requisiti, la cui origine, anche se non citata, non può che essere statistica, e le cui conseguenze sono esprimibili solo in termini statistici.

Più si procede nel progresso e più l'aspetto statistico invade e informa la Normativa.

Anche se si risale alle strategie di progetto delle strutture, cui si faceva ricorso alcuni decenni or sono (gli aspetti probabilistici della sicurezza

sono andati sempre via via evidenziandosi), si vede che i carichi, cui la struttura è effettivamente assoggettata in volo, vengono interpretati con valori F definiti probabilisticamente (ad es. sono superati in media se non una volta in un certo periodo di tempo). Allo stesso modo le resistenze (dei materiali e quindi delle strutture) vengono interpretati con valore R , probabili (cioè superati in pratica) al 99%, essendo tale stima credibile al 95%. Le resistenze R devono essere convenientemente superiori (di solito 1.5 volte) dei carichi F . Il complesso della strategia di sicurezza genera una situazione che in termini esatti potrebbe essere valutata solamente nei seguenti termini: sicurezza è «probabilità che un carico applicato sia inferiore alla resistenza della struttura». Il fatto che una tale valutazione sarebbe poco credibile a causa delle incertezze (attuali) sulle distribuzioni probabilistiche di carichi e resistenze soprattutto nelle zone lontane dai valori più probabili, nulla toglie all'aspetto eminentemente probabilistico della sicurezza che peraltro, nei termini predetti, nelle strutture aeronautiche risulta molto elevata (molto prossima a 1).

Oltre le strutture, di cui sono stati esposti alcuni punti connessi a particolari condizioni di lavoro, anche i vari componenti quali: motori, pompe, generatori, servocomandi, ecc. ecc., (tutti dispositivi che inevitabilmente sono soggetti a guasti, a malgrado dei controlli in costruzione e montaggio) sono fonti di rischi, per guasti singoli o combinati. Le strategie per la sicurezza comportano, sulla base della conoscenza delle probabilità di guasto di ogni componente, sostituzioni a tempo, duplicazioni, triplicazioni ecc. di individui in parallelo, ed altri accorgimenti, con lo scopo di ottenere probabilità di rischi anche qui estremamente remote.

È un dato di fatto che la «certezza» non è stata mai, fin qui, realizzata.

Ciò risulta dai dati disponibili. Anche esaminando velivolo per velivolo, causa per causa (struttura, motori, ambiente, fattori umani, ecc.) si trova una incidenza sulla percentuale di incidenti verificatisi.

Se si sposta l'analisi alla «possibilità» di ottenere *potenzialmente* la «certezza» nella sicurezza il risultato è univoco anche se esaminiamo separatamente i piani dottrinali e normativo. In entrambi si va via via consolidando la concezione probabilistica della sicurezza o, quanto meno, la derivazione probabilistica di molti requisiti o impostazioni nel progetto, che inevitabilmente fanno sì che il risultato globale sia di natura probabilistica.

L'approccio probabilistico allo studio del mezzo aereo iniziò nel periodo fra la 1^a e la 2^a Guerra Mondiale quando cominciarono a comparire gli aerei plurimotori. All'inizio l'analisi probabilistica fu utilizzata

unicamente per comparare fra loro le probabilità, relative alle varie soluzioni propulsive, di portare a termine un volo di lunga durata. Successivamente negli anni 30, quando le linee di trasporto aereo cominciarono ad espandersi, iniziarono ad essere elaborate statisticamente le avarie dei vari tipi di aerei riferite alle ore di volo per tipo di velivolo. Solamente nel 1939 venne quantizzato numericamente ad opera di Sir A.G. Pugsley quanto deve essere sicuro un aereo:

«il rateo di incidenti, considerando tutte le avarie possibili, deve essere inferiore a 10^{-5} per ora di volo e la quota parte di questa cifra dovuta ad avarie strutturali deve essere inferiore a 10^{-7} ».

Una presentazione organica dei concetti probabilistici sulla sicurezza è contenuta in [4].

In quest'opera si deve nettamente distinguere l'aspetto metodologico, che introduce alla visione probabilistica della sicurezza e rappresenta una testimonianza della natura del problema e delle strategie che la dottrina e la tecnica impiegano, dall'aspetto «filosofico» che presenta i principi e le valutazioni cui, secondo vari Autori, dovrebbe ispirarsi l'individuazione dei livelli accettabili di rischio, tenuto conto della natura del sistema.

Siamo qui più interessati al primo aspetto, che per noi è determinante nel tentativo di spiegare le implicazioni profonde della tecnica e delle connesse responsabilità. Per il secondo aspetto, non essendo in sede di discussione dottrinale, non c'è che rimettersi a quanto stabilito e fissato dalle normative vigenti e specificatamente a quelle relative al trasporto passeggeri.

Secondo Lievens, a fondamento del progetto, per quanto concerne la sicurezza, c'è la individuazione degli obiettivi di sicurezza.

«Avant d'entreprendre un programme d'élaboration et éventuellement de démonstration de la sécurité il importe de fixer les objectifs de ce programme le plus clairement possible».

Tale definizione passa attraverso più fasi componenti tra le quali, ad esempio:

- definizione del sistema,
- definizione della missione,
- definizione dell'impiego anormale.

In particolare la definizione della missione si compone di vari aspetti:

- specifiche del sistema,
- fasi delle missioni,

- ambiente,
- procedure di manutenzione e riparazione.

Fra questi *l'ambiente* può essere distinto in tipi: naturale, ridotto, comandato, artificiale, ecc.

In particolare si definisce:

«L'environnement induit résulte directement de la présence ou du fonctionnement du système (onde de choc de l'avion supersonique, gaz d'échappement de l'automobile, pollution par les usines...). Citons notamment:

les accélérations	les radiations nucléaires
l'explosion	la turbulence induite
le givrage	les vibrations mécaniques et acoustique
les chocs	la fumée
la température	l'électricité statique...»

«L'environnement naturel n'est pas affecté par l'intervention de l'homme ni par la présence du système; il n'est fonction que de l'espace et du temps».

Il inclut notamment:

les nuages	la neige
le brouillard	la pluie mêlée de neige
la pluie givrante	la foudre
la gelée	l'humidité
la grêle	les embruns salés
la pluie	le sable et la poussière
le vent et la turbulence	
la pollution atmosphérique	
les radiations solaires	
la température	
le magnétisme terrestre...	

È ormai invalso l'uso, e anche il Lievens lo fa, di introdurre una classificazione delle gravità e delle probabilità.

Classi di gravità

- | | | |
|-----------------------------|---|----------------------------------|
| — conseguenze minori | } | dette anche conseguenze maggiori |
| — conseguenze significative | | |
| — conseguenze critiche | | |
| — conseguenze catastrofiche | | |

Classi di probabilità

— avvenimenti frequenti	$P > 10^{-3}$;
— avvenimenti poco frequenti	$10^{-3} > P > 10^{-5}$;
— avvenimenti rari	$10^{-3} > P > 10^{-7}$;
— avvenimenti estremamente rari	$10^{-7} > P > 10^{-9}$;
— avvenimenti estremamente improbabili	$P < 10^{-9}$.

Il Lievens propone quindi i seguenti principi generali «obiettivi globali», (v. fig. 2.1):

- la comparsa di conseguenze catastrofiche deve essere estremamente rara o estremamente improbabile;
- la comparsa di conseguenze critiche deve essere rara o estremamente rara o estremamente improbabile.

Nous poserons comme principes généraux que:

- l'apparition de conséquences catastrophiques doit être extrêmement rare ou extrêmement improbable;
- l'apparition de conséquences critiques doit être rare, extrêmement rare ou extrêmement improbable.

Osservato poi, giustamente, che è spesso impossibile pervenire preventivamente ad una valutazione globale della probabilità in specie per le conseguenze critiche e catastrofiche, propone i seguenti criteri sostitutivi (obiettivi parziali):

- a) ogni guasto le cui conseguenze possono essere catastrofiche deve essere estremamente improbabile;
- b) ogni guasto le cui conseguenze possono essere critiche deve essere improbabile/leggi estremamente raro o estremamente improbabile;
- c) dispositivi di avvertimento o d'allarme devono essere previsti ogni volta che una situazione critica esige una azione di correzione immediata;
- d) ogni errore umano che possa avere conseguenze critiche deve essere improbabile o estremamente improbabile;
- e) ogni guasto le cui conseguenze possano essere significative non deve essere né frequente, né poco frequente.

«Nous verrons ultérieurement que, dans l'état actuel des connaissances, il est souvent impossible de parvenir à la probabilité globale d'apparition de conséquences catastrophiques (ou critiques). Des lors il faut avoir recours à des critères plus maniables; nous poserons que:

- a) toute panne dont les conséquences sont catastrophiques doit être extrêmement improbable;

- b) toute panne dont les conséquences sont critiques doit être improbable ou extrêmement improbable;
- c) des dispositifs d'avertissement ou d'alarme doivent être prévus a chaque fois qu'une situation critique exige une action correctrice immédiate;
- d) toute erreur humaine entraînant des conséquences critiques doit être improbable ou extrêmement improbable;
- e) une panne dont les conséquences sont significatives ne doit être fréquente ni peu fréquente».

Il citato A. osserva poi quanto segue:

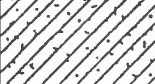
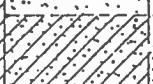
«Les critères a, b, c se retrouvent dans plusieurs textes réglementaires applicable (dans le présent ou dans le futur) au matériel aéronautique, notamment dans le TSS Standard 1.1, franco-britannique, applicable à Concorde; dans la condition CDQ*, française, applicable a l'Airbus, au Mercure, au Corvette, aux Falcon 10 et 50; dans les nouveaux amendements du 1309 de la FAR 25, américaine; dans le projet X Y Z relatif au futur règlement européen de navigabilité».

Ogni «funzione», sia essa necessaria o utile all'espletamento della missione, risulta soggetta ad una probabilità di non essere svolta o di non essere svolta correttamente, dal momento che ogni componente che vi concorre risulta soggetta ad una probabilità di rottura, guasto o errore.

Se per alcune funzioni (ad esempio la propulsione, i servocomandi, le comunicazioni ecc.) ciò risulta addirittura intuitivo, può risultare utile illustrare gli aspetti della funzione che meno di ogni altra a prima vista, può sembrare partecipe della predetta situazione ciò è: la funzione strutturale. Anzi, per tale funzione, giova distinguere le rotture «per fatica», le quali appaiono partecipi di una natura aleatoria a causa della aleatorietà delle condizioni di impiego, oltre che dal comportamento del materiale, sol che si abbia una modesta entrata nelle conoscenze tecniche-scientifiche coinvolte e nei dati sperimentali, dalle rotture per carichi «statici», cioè applicati «individualmente» e da paragonarsi alla «resistenza» della struttura stessa.

La figura 2.1 è stata riportata a puro titolo di esempio. Altre classificazioni e proposte esistono che però non mutano l'assunto centrale qui riportato.

Nelle pagine seguenti si propone una impostazione generale del problema della acquisizione di dati sulle probabilità di funzionamento dei componenti e di quello della individuazione della «probabilità di fun-

PROBABILITA' \ CONSEGUENZE	PROBABILITA'			
	10^{-5}	10^{-7}	10^{-9}	
	FREQUENTI O POCO FREQUENTI	RARI	ESTREMAMENTE RARI	ESTREMAMENTE IMPROBABILI
MINORI				
SIGNIFICATIVE				
CRITICHE				
CATASTROFICHE				

 GLOBALI,  PARZIALI

Figura 2.1 - Obiettivi di sicurezza secondo Lievens (valori da riferirsi alla attività aeronautica).

zionamento» di un sistema in un certo tempo e in certe condizioni, noti che siano i valori dei singoli componenti.

3. Fonti di dati sui componenti

Sono fonti di dati sulle probabilità di funzionamento o di guasto di componenti (v. fig. 3.1):

- 1) l'impiego di componenti identici, simili o assimilabili, (se l'impiego è fatto in laboratori si configura come «prova»);
- 2) la tecnologia attraverso lo studio dei processi di produzione e delle filosofie di progetto.

La dimostrazione di probabilità di guasto diviene più onerosa e incerta quanto più essa è remota.

Ogni dimostrazione di probabilità di guasto è soggetta a incertezza sulla probabilità che il campione esaminato sia interprete della popolazione.

Per i componenti standard il MIL HDBK-217 fornisce i tassi di guasto per differenti rapporti di utilizzazione e temperature.

Quale esempio di deduzione dei tassi di guasto dei dati storici si propone il caso di dati relativi a t , variabile casuale, di cui si possa supporre, ammetta una distribuzione esponenziale di densità

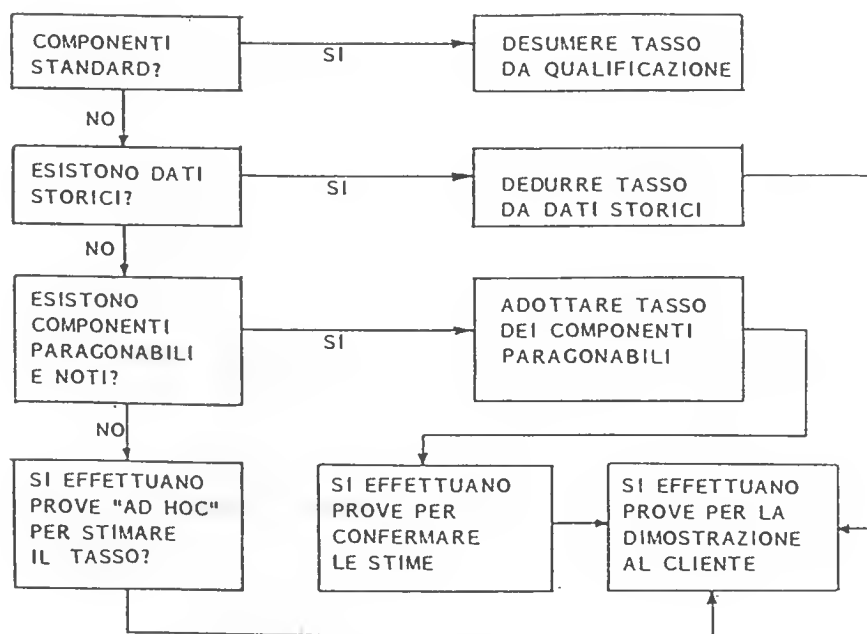


Figura 3.1 - Esempio di logica per la determinazione e la dimostrazione dei «tassi di guasto» (v. anche [12]).

$$f(t) = \lambda e^{-\lambda t}$$

in assenza di informazioni sulla distribuzione di λ .

Il problema che insorge è quello della probabilità che il campione, di dimensione n , interpreti la popolazione. Formiamo la media aritmetica

$$t_m^{(n)} = \frac{1}{n} (t_1 + \dots + t_n) .$$

Consideriamo la variabile casuale T , tale che

$$f(T) = \frac{1}{2} e^{-\frac{T^2}{2}} ,$$

formiamo

$$T_i^2: T_i^2 = T_{i,1}^2 + T_{i,2}^2 \quad (T_{i,j} \in T).$$

La variabile casuale

$$t_i = \frac{T_i^2}{2\lambda}$$

ammette una distribuzione esponenziale di densità.

$$f(t_i) = \lambda e^{-\lambda t_i}.$$

Quindi la variabile casuale $2\lambda n t_m^{(n)} = \zeta$, ha le espressioni

$$\zeta = T_1^2 + T_2^2 + \dots + T_n^2,$$

$$\zeta = T_{1,1}^2 + T_{1,2}^2 + T_{2,1}^2 + T_{2,2}^2 + \dots + T_{n,1}^2 + T_{n,2}^2.$$

La ζ è pertanto una variabile che ammette distribuzione χ^2 o di Pearson con $2n$ gradi di libertà.

Se χ_α^2 e $\chi_{1-\alpha}^2$ corrispondono a $R = 1 - \alpha$ e $R = \alpha$, si ha:

$$P(\chi_\alpha^2 < 2\lambda n t_m^{(n)} < \chi_{1-\alpha}^2) = 1 - 2\alpha,$$

$$P\left(\frac{2n\lambda t_m^{(n)}}{\chi_{1-\alpha}^2} < 1 < \frac{2n\lambda t_m^{(n)}}{\chi_\alpha^2}\right) = 1 - 2\alpha.$$

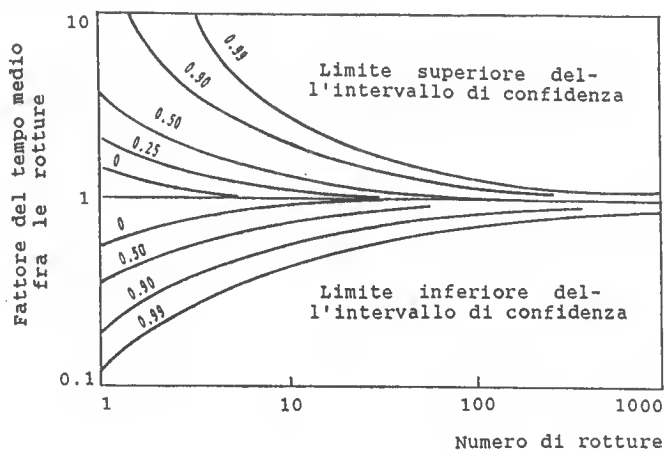
Definendo ora $1 - 2\alpha =$ livello di fiducia (o di confidenza), risulta

$$\frac{2n}{\chi_{1-\alpha}^2} \quad \text{-----} \quad \frac{2n}{\chi_\alpha^2} \quad \text{l'intervallo bilaterale di fiducia (o di confidenza)}$$

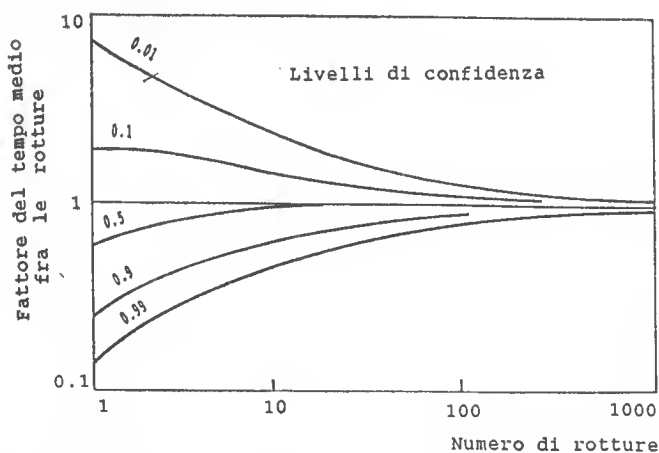
per $(t_m^{(n)} \cdot \lambda)^{-1}.$

In altre parole c'è probabilità $1 - 2\alpha$ (confidenza) che $(t^{(n)}\lambda)^{-1}$ sia contenuto nell'intervallo di confidenza. Si possono quindi tracciare curve

come quelle di fig. 3.2a-b, dove compaiono anche quelle relative a intervalli unilaterali.



(a)



(b)

Figura 3.2 - (a) Limiti di confidenza bilaterali per il tempo medio fra rotture (MTBF).

(b) Limite di confidenza inferiore unilaterale per il tempo medio fra rotture (MTBF).

In sintesi, se t è esponenziale, $t_m^{(n)}$ ha una distribuzione di Pearson di $2n$ gradi di libertà, ciò che consente di studiare livelli e intervalli di confidenza.

Nella fig. 3.3 sono sintetizzati alcuni risultati che si ottengono nel caso la popolazione ammetta altre distribuzioni o si desiderino livelli e intervalli di confidenza di grandezza varie.

Distribuzione	Quantità voluta	Noto	Incognita	Variabile stimatrice	
$N(\mu, \sigma^2)$	μ	σ	—	$u = \frac{\bar{u} - \mu}{\sigma/\sqrt{n}}$	$N(0, 1)$
	μ	—	σ	$t = \frac{\bar{x} - \mu}{s^2/\sqrt{n}}$	t_{n-1}
	σ	μ	—	$\eta = \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^2}$	χ_n^2
	σ	—	μ	$\eta = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\sigma^2}$	χ_{n-1}^2
	$\mu_1 - \mu_2$	—	$\sigma_1 = \sigma_2$		t
	$\mu_1 - \mu_2$	σ_1, σ_2	—		$N(0, 1)$
	$\sigma_1^2 - \sigma_2^2$	μ_1, μ_2	—		F
	$\sigma_1^2 - \sigma_2^2$	—	μ_1, μ_2		F
E	μ	—	—	$\eta = \frac{1}{n} \sum_{i=1}^n y_i$	χ_{2n}^2

Figura 3.3 - Variabili stimatrici e loro distribuzioni per alcune quantità volute.

Che $\zeta = 2\lambda n t_m^{(n)}$ ammette distribuzione χ^2 di Pearson, se t ammette distribuzione esponenziale, può essere dimostrato anche per altre vie.

Se per ipotesi

$$f(t) = \lambda c^{-\lambda t} \quad (t > 0),$$

definiamo

$$T = \sqrt{2\lambda t}, \quad T^2 = 2\lambda t.$$

Risulta

$$\left. \begin{aligned} \zeta = 2\lambda n t_m^{(n)} &= T_1^2 + T_2^2 + \dots + T_n^2 \\ &= 2\lambda t_1 + 2\lambda t_2 + \dots + 2\lambda t_n \end{aligned} \right\}.$$

Introduciamo ora

$$\zeta_i = a\lambda t_i \Rightarrow t_i = \frac{\zeta_i}{2\lambda}, \quad \frac{\partial t_i}{\partial \zeta_i} = \frac{1}{2} \lambda^{-1}.$$

Risulta

$$f(\xi_i) = \lambda e^{-\frac{\zeta_i}{2}} \cdot \frac{1}{2\lambda} = \frac{1}{2} e^{-\frac{\zeta_i}{2}}$$

la variabile ζ è anche ottenibile come

$$\zeta = \sum_{i=1}^n \zeta_i.$$

La d.d.p. della ζ può essere ottenuta attraverso la funzione caratteristica. Questa coincide con il valore medio della variabile $e^{jp\zeta}$ e viene espressa dalla trasformazione di Fourier della $f(\zeta)$

$$E(e^{jp\zeta}) = \int_{-\infty}^{\infty} e^{jp\zeta} f(\zeta) d\zeta,$$

si che la $f(\zeta)$, da noi voluta, può essere espressa come la antitrasformata di $e^{jp\zeta}$. Poiché siamo in grado di calcolare per altra via $E(e^{jp\zeta_i})$, possiamo procedere nel modo seguente

$$E(e^{jp\zeta}) = E(e^{jp\sum \zeta_i}) = E\left(\prod_{i=1}^n e^{jp\zeta_i}\right) = \prod_{i=1}^n [E(e^{jp\zeta_i})].$$

D'altra parte è

$$E(e^{ip\xi_i}) = \frac{1}{2} \int_0^\infty e^{ip\xi_i} e^{-\frac{1}{2}\xi_i} d\xi_i = \frac{1}{2} \int_0^\infty e^{\left(-\frac{1}{2}+ip\right)\xi_i} d\xi_i = \frac{1}{1-2_j p}.$$

Quindi risulta in definitiva:

$$E(e^{ip\xi}) = \left(\frac{1}{1-2_j p} \right)^n.$$

Applicando ora l'antitrasformata:

$$f(\xi) = \frac{1}{2\pi} \int_{-\infty}^\infty \left(\frac{1}{1-2_j p} \right)^n e^{-ij\xi} dp = \frac{1}{2^n(n-1)!} \xi^{n-1} e^{-\frac{n}{2}\xi},$$

che è la densità di probabilità della variabile casuale χ^2 di Pearson, con $2n$ gradi di libertà.

Più sofisticate considerazioni devono essere fatte quando i dati disponibili si innestano in una esperienza precedente di cui occorre tenere conto. Il teorema di Bayes della probabilità condizionata acquista in questi casi una posizione centrale, dando luogo alle così dette «tecniche Bayesiane» di correzione. A titolo di esempio si riporta qui il caso in cui si possa ritenere dall'esperienza che il tasso di guasto λ ammetta distribuzione $\Gamma(\lambda)$ di media $\bar{\lambda}$ e scarto q in σ , e siano disponibili, per un singolo componente i dati del numero di guasti k in un tempo r . La correzione segue il metodo di Bayes, e consiste nella determinazione del valore corretto della media $\hat{\lambda}$.

Formati i parametri $\alpha = \frac{\bar{\lambda}}{\sigma^2}$ e $\beta = \frac{\lambda^2}{\sigma^2}$, la distribuzione $\Gamma(\lambda)$ ammette densità:

$$f(\lambda) = \frac{\lambda^{\beta-1} e^{-\alpha\lambda}}{\alpha^{-\beta} \Gamma(\beta)}.$$

Con riferimento alla trattazione precedente è ora $n=R$; $2\lambda nt_{(n)} = 2\lambda\tau$. Quest'ultima variabile se t è esponenziale, ammette distribuzione di Pearson con $2r$ gradi di libertà. Pertanto la probabilità di aversi una media aritmetica $2\lambda\tau$, condizionate dall'essere il tasso pari a $\bar{\lambda}$, è

$$P(2\lambda\tau/\bar{\lambda}) = \int_0^\infty \left(\frac{1}{2^k (k-1)!} \right) (2\bar{\lambda}\tau)^{k-1} e^{-\bar{\lambda}\tau} d\lambda.$$

La correzione Bayesiana, nella concezione di Robbins, esprime $\hat{\bar{\lambda}}$, nel modo seguente:

$$\hat{\bar{\lambda}} = \frac{\int_0^\infty \left(\frac{\lambda}{2^k (k-1)!} \right) (2\bar{\lambda}\tau)^{k-1} e^{-\bar{\lambda}\tau} \bar{\lambda} \left(\frac{\bar{\lambda}^{\beta-1} e^{-\lambda\alpha}}{\alpha^{-\beta} \Gamma(\beta)} \right) d\bar{\lambda}}{\int_0^\infty \frac{\lambda}{2^k (k-1)!} (2\bar{\lambda}\tau)^{k-1} e^{-\bar{\lambda}\tau} \frac{\lambda^{\beta-1} e^{-\alpha\bar{\lambda}}}{\alpha^{-\beta} \Gamma(\beta)} d\bar{\lambda}}$$

tenuto anche conto dello Jacobiano $\alpha^{-\beta}$ del cambiamento di variabile. Questa equazione in $\bar{\lambda}$, dà

$$\hat{\bar{\lambda}} = \frac{k + \beta}{\tau + \alpha}.$$

Analogamente, volendo la correzione $\hat{\sigma}d\sigma$, la si può ottenere da una formula che usi il momento secondo, rispetto alla media, che sostituisca cioè al numeratore nell'integrando

$$\left(\bar{\lambda} - \frac{k + \beta}{\tau + \alpha} \right)^2 \quad \text{a} \quad \bar{\lambda}.$$

Il risultato è

$$\hat{\sigma}^2 = \frac{k + \beta}{(\tau + \alpha)^2}.$$

4. Valutazione complessiva delle probabilità

In questo paragrafo si propone una interpretazione del tutto generale del problema della valutazione complessiva della probabilità di funzionamento di un sistema. Tale interpretazione è fondata su due noti teoremi:

- a) la probabilità che si guastino entrambi due sistemi è il prodotto delle probabilità che si guasti ciascuno dei due;
- b) la probabilità che si guasti uno o l'altro di due sistemi, è la somma delle probabilità di ciascuno, meno il prodotto delle due probabilità.

Come è noto vi sono varie distribuzioni di probabilità (esponenziale, gaussiana, binomiale, etc.), che possono adattarsi ciascuna alla natura di un componente o di un evento. Casi estremi sono quelli degli eventi discreti per i quali possono valere distribuzioni espresse da funzioni generalizzate.

I due teoremi suddetti valgono indipendentemente dalle distribuzioni. La fig. 4.1 contiene la dimostrazione del teorema b).

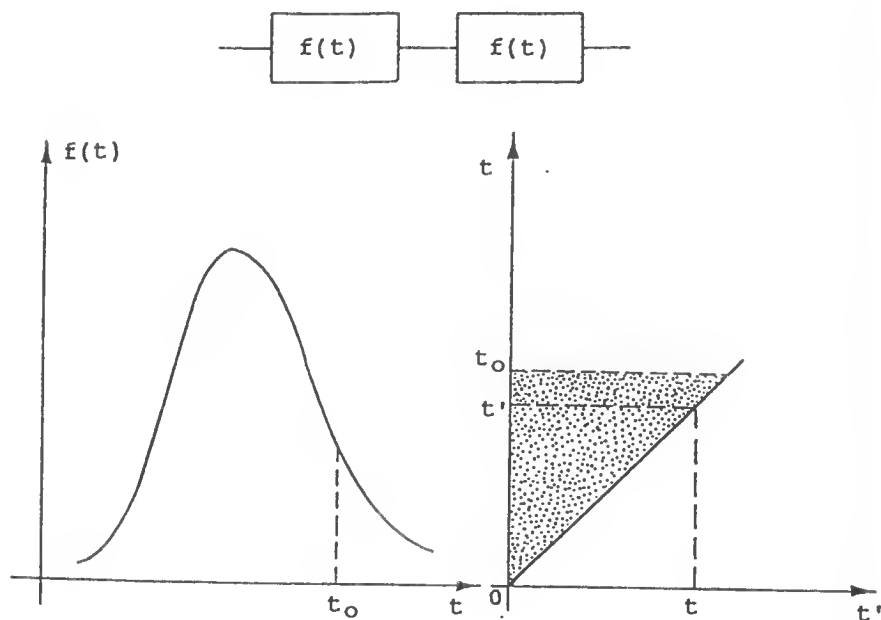


Figura 4.1 - Probabilità di guasto in t_0 .

$$1) \quad F = 2 \int_0^{t_0} f(t) dt - 2 \int_0^{t_0} f(t) dt \int_0^t f(t') dt' .$$

Poiché

$$\int_0^{t_0} dt \int_0^t f(t) f(t') dt = \int_0^{t_0} dt \int_t^{t_0} f(t) f(t') dt ,$$

si ha anche:

$$2) \quad F = 2 \int_0^{t_0} f(t) dt - 2 \int_0^{t_0} f(t) dt \int_t^{t_0} f(t') dt' .$$

Sommando 1) e 2), si ottiene in definitiva:

$$F = 2 \int_0^{t_0} f(t) dt - \int_0^{t_0} f(t) dt \int_0^{t_0} f(t) dt' .$$

Sia m il numero dei componenti, comprese le azioni umane, che fanno parte di un sistema nel suo complesso.

Sia p_i la probabilità di funzionamento dell' i -esimo componente in un tempo prestabilito ($i = 1 \dots m$).

Il numero n di eventi semplici, fra loro indipendenti, quando per ciascun elemento si considerino, nel tempo prestabilito due possibilità: funzionamento e guasto, è

$$n = 2^m .$$

In ciascun *evento semplice* vi sono componenti in funzione (siano f) e componenti guasti ($m-f$), e l'e.s. ha probabilità:

$$P_e \doteq p_1 p_2 \dots p_f (1-p)_{f+1} \dots (1-p)_m$$

$$e = 1, \dots, n .$$

Vale ovviamente la relazione:

$$\sum_{e=1}^n P_e = 1 .$$

La probabilità di funzionamento P_s è la somma (≤ 1) delle probabi-

lità $P_{e,f}$ degli eventi che assicurano il funzionamento

$$P_s = \sum P_{e,f}.$$

Con una impostazione siffatta è possibile, di principio, risalendo agli eventi semplici indipendenti, ottenere la probabilità di funzionamento di un sistema comunque complesso e condizionato.

In molti casi, per economia di tempo e di mezzi, si procede per vie semplificate, le quali, correttamente da un punto di vista ingegneristico, valutano maggioranti o minoranti dei valori esatti, sempre ponendo come condizione per l'accettazione della approssimazione, che la sicurezza o l'affidabilità o qualsiasi altra caratteristica in esame sia in effetti migliore o superiore a quella calcolata.

Molte delle considerazioni che seguono sono rivolte a introdurre e analizzare modi e criteri per approssimazioni di tale natura.

Una rappresentazione grafica potrebbe essere fatta avvalendosi di spazi a m dimensioni oppure con il corrispondente diagramma di Weiss (fig. 4.2).

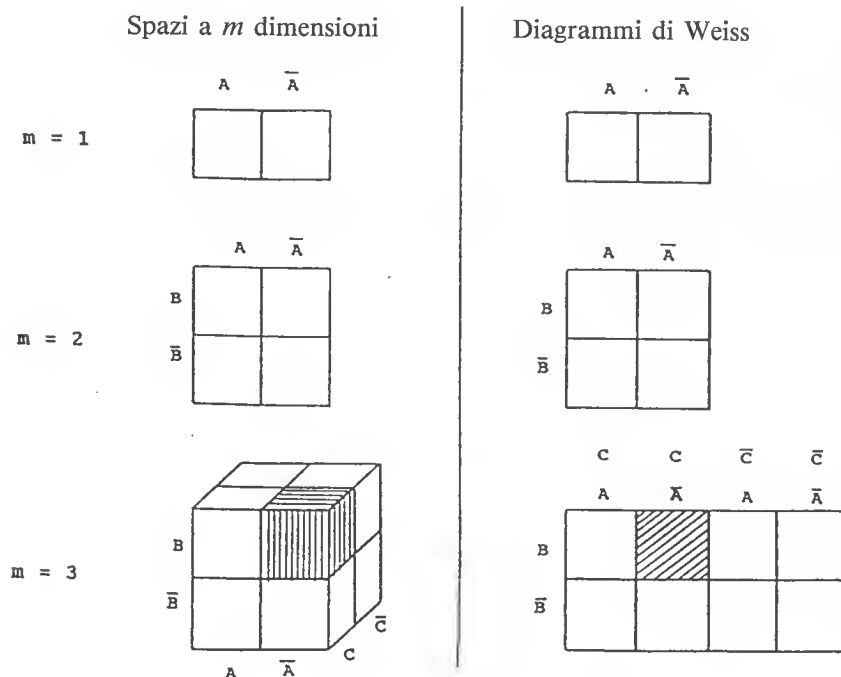


Figura 4.2 - Spazi a m dimensioni e diagrammi di Weiss.

La somma delle probabilità di ogni suddivisione vale uno.

Quando si usa una rappresentazione del genere il problema è quello di fare una partizione delle suddivisione in due parti:

F = insieme di tutte le suddivisioni che assicurano il funzionamento

$\bar{F}$ = insieme di tutte le suddivisioni che danno luogo a non funzionamento.

Per fare in modo esatto tale partizione vi sono sostanzialmente due modi di procedere, l'analisi con modi di funzionamento e l'analisi con tagli di guasti.

Un esempio elementare di analisi con modi di funzionamento è quello di fig. 4.3. Individuati i nodi 1, 2, 3 e 4 dello schema di funzionamento, si ricava la matrice di funzionamento con connessioni del 1° ordine:

CONNESSIONI 1° ORDINE

	1	2	3	4
1	0	A	B	0
2	0	0	0	C
3	0	0	0	C
4	0	0	0	0

Il prodotto di tale matrice per se stessa genera quella delle connessioni del 2° ordine:

CONNESSIONI 2° ORDINE

1	0	0	0	AC + BC
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0

Non v'è luogo a considerare connessioni del 3° ordine. Il funzionamento, cioè la connessione da 1 a 4, si può esprimere con

$$F = AC + BC.$$

Operando con algebra di Boole, si ha

$$F = AC(B + \bar{B}) + BC(A + \bar{A})$$

$$F = ACB + AC\bar{B} + ABC + CB\bar{A}.$$

Nella figura, le suddivisioni che assicurano il funzionamento sono tratteggiate, mentre la sopralineatura significa *non funzionamento*.

Sullo stesso sistema di fig. 4.3 una analisi con tagli di guasti, individua il non funzionamento con il verificarsi delle due condizioni $\bar{A}\bar{D}$ e $\bar{C}$

$$N.F. = \bar{A}\bar{B} + \bar{C}.$$

Operando con l'algebra di Boole, si ha

$$N.F. = \bar{A}\bar{B}(C + \bar{C}) + \bar{C}(A + \bar{A})(B + \bar{B})$$

$$N.F. = \bar{A}\bar{B}\bar{C} = \bar{A}\bar{B}C + \bar{A}\bar{B}\bar{C} + \bar{A}\bar{C}B + \bar{A}\bar{C}\bar{B} + \bar{A}\bar{B}\bar{C}$$

è facile verificare che

$$F + N.F. = 1$$

come è necessario.

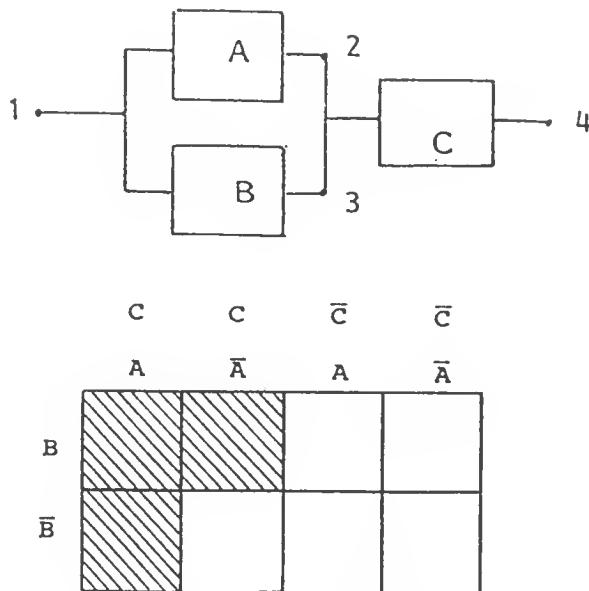


Figura 4.3.

Sullo stesso sistema si può applicare un esempio elementare di valutazioni approssimate, (fig. 4.4).

Le seguenti condizioni di guasto:

$$\begin{cases} \bar{A}, \bar{B} \\ \bar{C} \end{cases} \quad \text{di probabilità} \quad \begin{cases} (1 - R_A)(1 - R_B) \\ (1 - R_C) \end{cases}$$

sono fra loro indipendenti (un guasto entra in una sola condizione). Non è possibile individuarne altre.

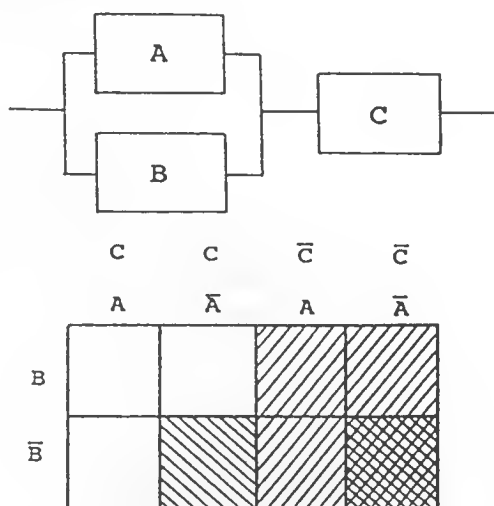


Figura 4.4 - Valutazioni approssimate.

La somma delle probabilità delle condizioni di guasto è $\geq$ alla probabilità esatta di guasto.

$$P(\bar{A}, \bar{B}) + P(\bar{C}) > P(\bar{A}\bar{B}\bar{C}) + P(\bar{A}\bar{B}C) + P(\bar{A}\bar{C}B) + P(A\bar{B}\bar{C}) + P(A\bar{B}C).$$

In questo caso l'approssimazione consiste nel conteggiare due volte la suddivisione $\bar{B}\bar{C}\bar{A}$ tra le cause di non funzionamento.

5. Logiche decisionali per la sicurezza e gli altri requisiti

I principali passi per il «progetto» della sicurezza, dell'affidabilità e dei concetti collegati sono:

- analisi dei rischi (a fronte della missione);
- analisi della sicurezza e dell'affidabilità;
- analisi dei modi di rottura e delle conseguenze;
- analisi della disponibilità (e dei concetti collegati);
- analisi della manutenibilità (se necessaria).

L'analisi dei rischi consiste nell'identificare i «rischi potenziali» da considerare nelle analisi di sicurezza e affidabilità.

In questa analisi vanno comprese tutte le eventualità create in qualsiasi modo, ivi comprese le conseguenze di:

- effetti di rotture sui componenti vicini;
- effetti di errori di installazione;
- incendi e esplosioni;
- errori umani;
- etc.

Identificati i rischi, si attua nei loro confronti una strategia, che in modo semplificato è condensata in fig. 5.1. La semplificazione principalmente è connessa al fatto che vi sono casi nei quali si hanno «allarmi» e «emergenza» anche in presenza di dispositivi di sicurezza.

L'analisi della sicurezza e dell'affidabilità comporta:

- definire gli obiettivi di sicurezza e affidabilità a livello di sistema (di sottosistema);
- validare le scelte di «architettura» capaci di soddisfare gli obiettivi, tenendo conto delle interfacce a livello di sistema (di sottosistema);
- specificare i requisiti di sicurezza e affidabilità per i sottosistemi (per i componenti);
- valutare sicurezza e affidabilità del complesso integrando i dati dei sottosistemi (dei componenti).

L'analisi dei modi di rottura e delle conseguenze si fonda sulla individuazione delle «interruzioni di funzioni» e nella loro classificazione in funzione della gravità delle conseguenze.

Ad esempio si possono prendere in conto i seguenti livelli di gravità:

- minori, maggiori, pericolose e catastrofiche.

Quali requisiti generali di sicurezza e affidabilità si pongono obiettivi che, in base alla loro natura, possono essere classificati in: «obiettivi qualitativi» e «obiettivi quantitativi».

Quale esempio di prescrizioni qualitative possiamo fare riferimento ai requisiti di aversi «operatività nella rottura» dopo una interruzione

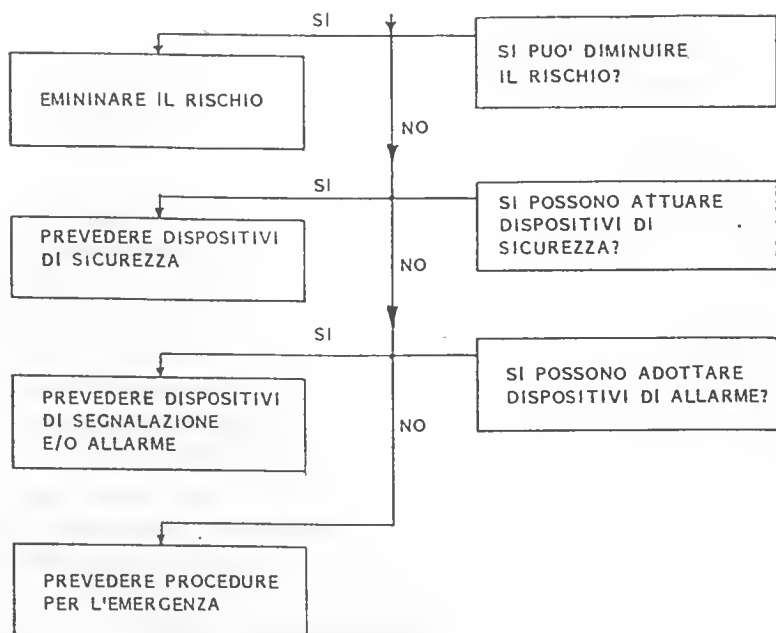


Figura 5.1 - Logica decisionale semplificata per la sicurezza (v. anche [12]).

Nota: Tra i dispositivi di sicurezza vanno inclusi i «margini» da assumersi in progetto.

«maggiore» e «sicurezza nella rottura» dopo una interruzione «rischiosa» o una «catastrofica».

Quale esempio di prescrizioni quantitative possiamo invece fare riferimento ai requisiti di aversi:

per una «interruzione maggiore» probabilità $< 10E - 4$,

per una «interruzione rischiosa» probabilità $< 10E - 5$,

per una «interruzione catastrofica» probabilità $< 10E - 6$,

oltre a prescrizioni relative all'intero sistema.

Nella attività spaziale, dove recentemente questi concetti hanno ricevuto un notevole impulso, si registra un significativo intreccio tra le vicende e il pensiero di progetto.

Alle origini la NASA nel 1959 scelse il lanciatore Redstone invece di Atlas in base alle probabilità di successo, anche se Atlas aveva maggiori capacità in quanto a carichi utili.

Redstone aveva solo capacità sub-orbitali. Atlas però aveva dimostrato solo il 40% di probabilità di successo. Così come Thor, Vanguard e Titan a quello stadio di sviluppo.

La NASA capì in quella circostanza:

- 1) che l'attività spaziale è «a rischio»;
- 2) che è necessario un «risk assessment» attraverso la storia operativa.

Pose infatti il limite di 80% per l'obiettivo di affidabilità (missioni senza equipaggi) di «safe orbit».

Il programma Apollo, successivamente, fu fondato su un criterio quantitativo dello 0.95% di probabilità di sicurezza per la missione e dello 0.995 per l'equipaggio. Dopo 17 missioni (un solo insuccesso) risultava poi dimostrata una probabilità di successo delle missioni dello 0.94.

Oggigiorno la situazione è molto differente dalle origini. V'è una richiesta dal pubblico e una confidenza dei progettisti verso sicurezze più elevate. Ne è prova il fatto che una perdita di missione (51L) ha messo in forse l'intero programma NASA. D'altra parte i successi precedenti hanno creato l'idea di «attenersi a ciò che ha funzionato», idea rafforzata anche dalla scoperta che alcune previsioni risultavano irrealistiche.

Tutto ciò portò la NASA ad un mutamento fondamentale di atteggiamento. Da una impostazione quantitativa a una qualitativa, esemplificata dalla metodologia FMEA-CIL.

Il criterio quantitativo necessita di stabilire credibili «data base» attraverso la raccolta di dati e la loro analisi.

Il discredito sul «quantitativo» fece sì che non si sviluppassero i «data base» come in altre tecnologie.

Dopo l'incidente 51L, la NASA ha iniziato studi che hanno raccolto dati, di quantità e natura preliminare, ma tali da far pensare che fosse possibile, da un punto di vista tecnico statistico, creare credibili «data base», continuando su quella strada.

Alla creazione di «data base» si oppongono tuttavia, se non propriamente degli ostacoli, delle difficoltà, originate da motivi storici, tecnologici, brevettuali, finanziari, di sicurezza e istituzionali.

I vari concetti che fanno ormai parte integrante del progetto, sono a volte fra loro in opposizione, sì da giustificare soluzioni diverse per missioni differenti.

Ad esempio, in un sistema «manutenibile» la ridondanza contrasta con la «maintainability» e anche con l'«availability».

In sistemi come le «space stations», per citare un caso, i dati (quantitativi) di «reliability» sono fondamentali. In mancanza i requisiti di «maintainability e supportability» non possono essere definiti e le operatività della stazione è sospetta.

Nella missione Apollo si ha missione unica (o breve) quindi i requisiti di successo sono ottenuti in progetto e non necessita la maintainability, quindi reliability e availability sono sinonimi.

Nella missione Shuttle che è plurimissione (breve) l'availability principale è quella al «lancio» e la maintainability è di poca importanza e riguarda la prontezza operativa; quindi reliability e availability sono distinte.

Infine, nella missione Freedom (Space Station) avente durata operativa di 30 anni la ridondanza ha effetti importanti e la maintainability e la supportability sono decisive; le varie reliability, availability, maintainability e supportability sono distinte.

La necessità di provvedere a predisporre «maintainability» e «supportability» rende fondamentale la capacità di dominare *qualitativamente* la reliability.

Una convincente filosofia di progetto deve prevedere prescrizioni di natura qualitativa e quantitativa combinate.

Infatti prescrivere il soddisfacimento del criterio «Fail-Safe» o «Fail-Operativo» non dà alcuna garanzia effettiva se non si accompagna il criterio stesso con un requisito sulla probabilità del guasto.

D'altra parte, un requisito quantitativo che imponga alla probability di essere molto piccola, potrà essere ottenuto credibilmente solo con accorgimenti di duplicazione, triplicazione, etc. di dispositivi di cui si disponga di valutazioni delle probabilità di guasto, «credibili» (molto probabilmente inferiori ai requisiti).

6. Metodologia automatiche per i sistemi

6.1 Metodo di Montecarlo

Fra le metodologie automatiche, da utilizzarsi per la stima delle caratteristiche di affidabilità e/o sicurezza di sistemi, un posto particolare spetta alla simulazione, in particolare quella nota come «metodo Montecarlo», che, oltre simulare il comportamento del sistema, mediante il suo modello matematico, simula anche la casualità degli eventi, grazie alle generazioni di numeri casuali, prestazione oggi disponibile da software su praticamente ogni tipo di elaboratore elettronico. In questo modo si può simulare la casualità dell'occorrenza dei guasti, su un componente (fig. 6.1) di tasso di guasto costante nel tempo « λ », in un tempo « t »; la stima, rifacendosi ad un «esperimento» (seppure simulato in luogo di essere fisicamente effettuato) diventa probante solo dopo un certo numero di simulazioni effettuate, assumendo come risultato la media dei risultati delle diverse simulazioni.

La simulazione può essere applicata a sistemi di qualsiasi complessità (v. ad esempio fig. 6.2). In tal caso, per ogni missione simulata occorrono tanti numeri casuali quanti sono i componenti.

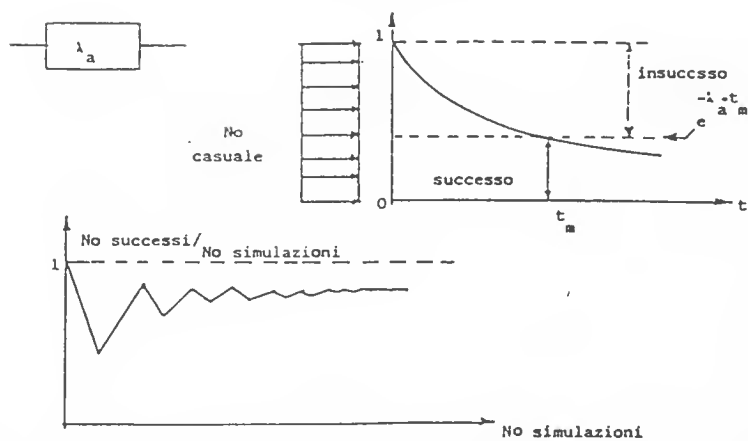


Figura 6.1 - Metodo di Montecarlo simulazione con numeri casuali, [12].

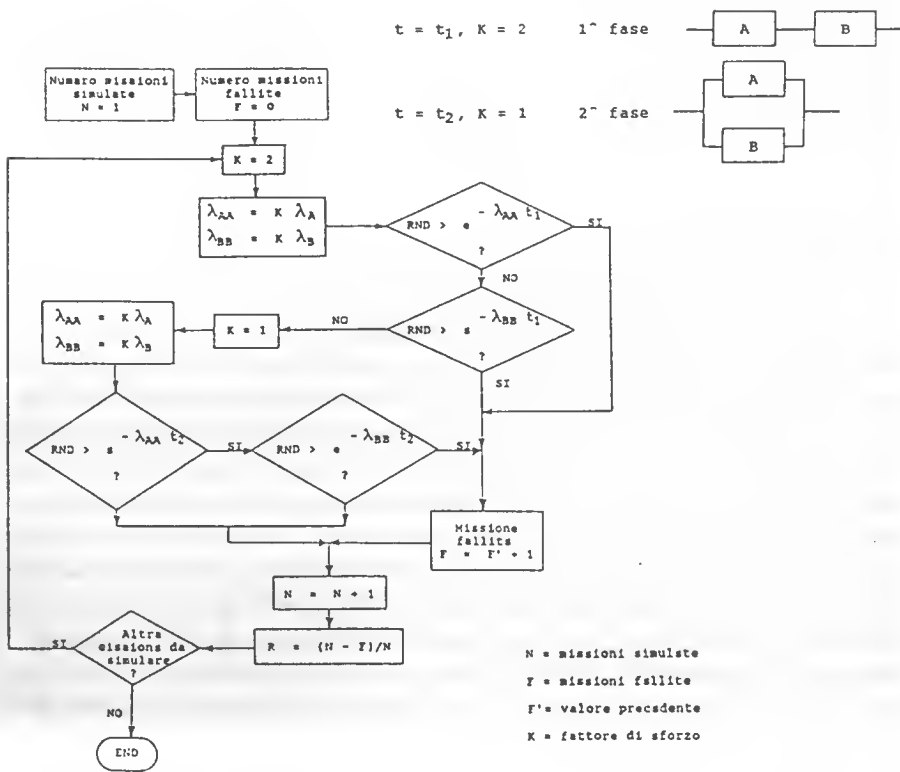


Figura 6.2 - Esempio di simulazione con metodo di Montecarlo, [12].

6.2 Programmazione dinamica

È una metodologia automatica che viene in particolare usata per la allocazione di probabilità di guasti.

È applicabile ai problemi decisionali (o sequenziali) multistadio, convertendoli in problemi di ottimizzazione monostadio.

Occorre definire gli stadi del processo e in particolare identificare (v. fig. 6.3) [11]:

- S_k , l'insieme dei possibili stati $s_k \in S_k$ dello stadio k ;
- D_k , l'insieme delle possibili decisioni $d_k \in D_k$ alternative disponibili allo stadio k ;
- $T_k(s_k, d_k)$, la funzione che trasforma s_k in s_{k-1} ;
- $R_k(s_k, d_k)$, la funzione che definisce il ritorno realizzato allo stadio k in conseguenza dello stato s_k e dell'alternativa d_k .

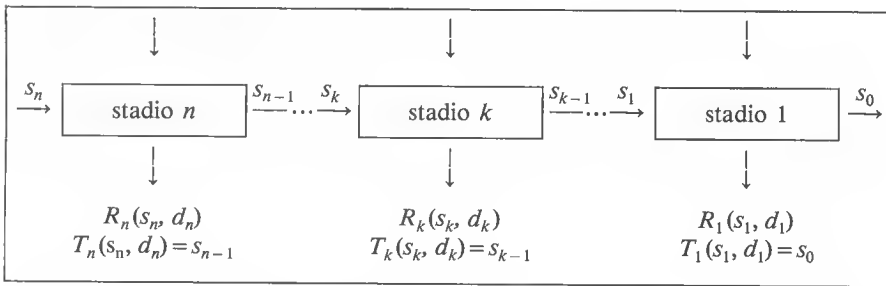


Figura 6.3 - Schema di processo multistadio.

Il problema può essere posto sotto forma di equazioni ricorrenti

$$f_k(s_k) = \min_{d_k \in D_k} [Q_k(s_k, d_k)] , \quad k = 1, 2, \dots, n$$

ove

$$\begin{aligned} Q_k(s_k, d_k) &= R_k(s_k, d_k), & k &= 1 \\ &= R_k(s_k, d_k) \cdot f_{k-1}(s_{k-1}), & k &= 2, 3, \dots, n \end{aligned}$$

dove il segno $\cdot$ può rappresentare somma o prodotto (in caso di operandi non negativi).

Si ha il ritorno totale, associato alla scelta ottima di alternative d^* :

$$f_n(s_n) = \min_{d_n \in D_n} [R_n(s_n, d_n) \cdot f_{n-1}(s_{n-1})],$$

ove $d^* = (d_1^*, d_2^*, d_3^* \dots d_n^*)$.

Sostituendo a $f_{n-1}(s_{n-1})$ la sua espressione ricorrente, si ottiene in definitiva:

$$f_n(s_n) = \min_{d_n \in D_n} \left\{ R_n(s_n, d_n) + \min_{d_{n-1} \in D_{n-1}} R_{n-1}(s_{n-1}, d_{n-1}) + \left[\min_{d_{n-2} \in D_{n-2}} R_{n-2}(s_{n-2}, d_{n-2}) + \dots + \left(\min_{d_1 \in D_1} R_1(s_1, d_1) \right) \right] \right\}.$$

La programmazione dinamica è applicabile alle serie (logiche) di sottosistemi, per minimizzare lo sforzo complessivo di miglioramento delle probabilità di funzionamento fino a 1 con le assunzioni seguenti ($0 \leq \bar{y} \leq 1$).

I) Il sistema è composto da n sottosistemi in serie dei quali sono «note» le probabilità ($0 \leq x_i \leq 1$) attuali di funzionamento ed esse non bastano allo scopo di ottenere un obiettivo $\bar{y}$ di probabilità totale ($0 \leq \bar{y} \leq 1$) e dovranno essere aumentate fino alle y_i . Nessuna probabilità può essere diminuita ($x_i \leq y_i \leq 1$).

II) Gli n sottosistemi sono indipendenti. Quindi la probabilità complessiva è:

$$y = \prod_{i=1}^n y_i$$

se y_i è la probabilità dell' i -esimo componente.

III) Può essere individuato uno sforzo ($G_i(x_i, y_i)$) necessario per ogni componente per innalzarne la probabilità dal valore attuale a un valore qualsiasi variabile discretamente o con continuità).

IV) Le probabilità della soluzione ottima sono le y_i^* .

Il problema può essere formulato nel modo seguente:

$$\begin{aligned} \min \quad & \sum_{i=1}^n G_i(x_i, y_i) \\ \text{con} \quad & \prod_{i=1}^n y_i = \bar{y} . \end{aligned}$$

Il problema stesso può essere convertito in programmazione dinamica con le seguenti posizioni:

- 1) ogni sottosistema è assimilato a uno stadio per il quale si può individuare una «allocazione obiettivo».
- 2) Si definisce S_k in modo che

$$1 = s_n \geq s_{n-1} \dots \geq s_1 \geq s_0 = \bar{y} .$$

- 3) Si definisce D_k di ogni possibile alternativa con $d_k = y_k$, in modo che

$$x_k \leq y_k \leq 1 \quad (k = 1, 2, \dots, n) .$$

- 4) Si definisce $T_k(s_k, d_k)$ nel modo seguente:

$$s_k y_k = s_{k-1} .$$

- 5) Si definisce

$$R_k(s_k, d_k) = G_k(x_k, y_k) .$$

Ne risultano le equazioni ricorrenti:

$$f_k(s_k) = \min_{y_k} [Q_k(s_k, y_k)]$$

$$\begin{aligned} Q_k(s_k, y_k) &= G_k(s_k, y_k) , & k &= 1 \\ &= G_k(s_k, y_k) + f_{k-1}(s_{k-1}) , & k &= 2, 3, \dots, n . \end{aligned}$$

6.3 Processi di Markov

Un sistema nel quale i componenti sono soggetti a guasti e possono essere riparati, si che quanto al tempo medio fra le rotture si abbia un tempo medio per le riparazioni, può essere analizzato per mezzo dei processi marcoviani. Solitamente rotture e riparazioni si ammette abbiano distribuzioni esponenziali.

La fig. 6.4 presenta un esempio di applicazione di un processo discreto (catena) di Markov a un sistema composto dal parallelo di A e B, [15].

I possibili stati (cioè gli eventi semplici del para. 4) sono 4, condensabili in 3, perchè $\bar{A}B$ e $A\bar{B}$ danno gli stessi effetti e se ne può considerare l'unione. Ciascuno degli stati ha una probabilità di verificarsi ad un dato tempo t . Nel successivo intervallo t vi è una certa probabilità di passaggio da ciascuno stato a ciascuno dei tre considerati. Le probabilità in esame sono quindi 9 e possono costituire una matrice S (di Markov), tale che se $\mathbf{P}(t)$ è la colonna ordinata di probabilità di trovarsi negli stati 0, 1 e 2 al tempo t , la corrispondente colonna al tempo $t + Dt$ è

$$\mathbf{P}(t + Dt) = S \mathbf{P}(t).$$

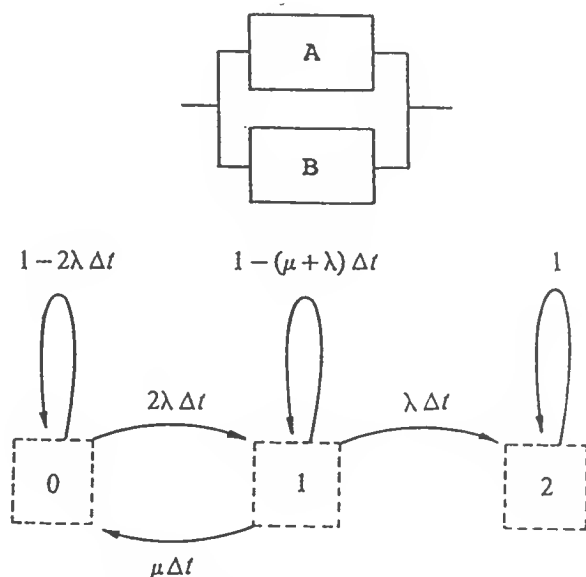


Figura 6.4 - Approccio alla Markov di un sistema con riparazioni.

Nel caso più generale questa relazione, purché P sia una colonna comprendente tutti gli stati, può essere usata per qualsiasi sistema. Se λ = tasso

di guasto di A e B , μ = tasso di riparazione di A e B , l'approccio alla Markov dà:

$$P(t) \begin{cases} P_0(t) \\ P_1(t) \\ P_2(t) \end{cases} \quad (P_0 + P_1 + P_2 = 1)$$

$$P(t + \Delta t) \begin{cases} P_0(t + \Delta t) = P_0(t) (1 - 2\lambda \Delta t) + P_1(t) \mu \Delta t \\ P_1(t + \Delta t) = P_0(t) 2\lambda \Delta t + P_1(t) [1 - (\mu + \lambda) \Delta t] \\ P_2(t + \Delta t) = P_1(t) \lambda \Delta t + P_2(t) \end{cases} .$$

Indicando con $\dot{P}_i(t)$ i cambiamenti di probabilità di ciascuno stato i -esimo, nel tempo Δt , si ha:

$$\dot{P}(t) \begin{cases} \dot{P}_0(t) = -2\lambda P_0(t) + \mu P_1(t) , \\ \dot{P}_1(t) = -(\mu + \lambda) P_1(t) + 2\lambda P_0(t) , \\ \dot{P}_2(t) = \lambda P_1(t) ; \end{cases}$$

ovvero:

$$\dot{P}(t) = M P(t) \Delta t; \quad M = \begin{vmatrix} -2\lambda & \mu & 0 \\ 2\lambda & -(\mu + \lambda) & 0 \\ 0 & \lambda & 0 \end{vmatrix} .$$

Valgono pertanto, anche nel caso più generale, le seguenti relazioni, che prevedono molti passi Δt :

$$\begin{cases} P(\Delta t) = P(0) + M P(0) \Delta t \\ \dots \\ P(n \Delta t) = P[(n-1) \Delta t] + M P[(n-1) \Delta t] \Delta t = \\ = (I + \Delta t M) P[(n-1) \Delta t] = S P[(n-1) \Delta t] ; \end{cases}$$

dove

$$P(n \Delta t) = S^n P(0) \quad \text{e} \quad S = I + M \Delta t .$$

A livello di impiego pratico è stato proposto [15] di sveltire i tempi di calcolo, osservando che S_n può essere ottenuto con un limitato numero di moltiplicazioni p , sfruttando uno schema del tipo seguente:

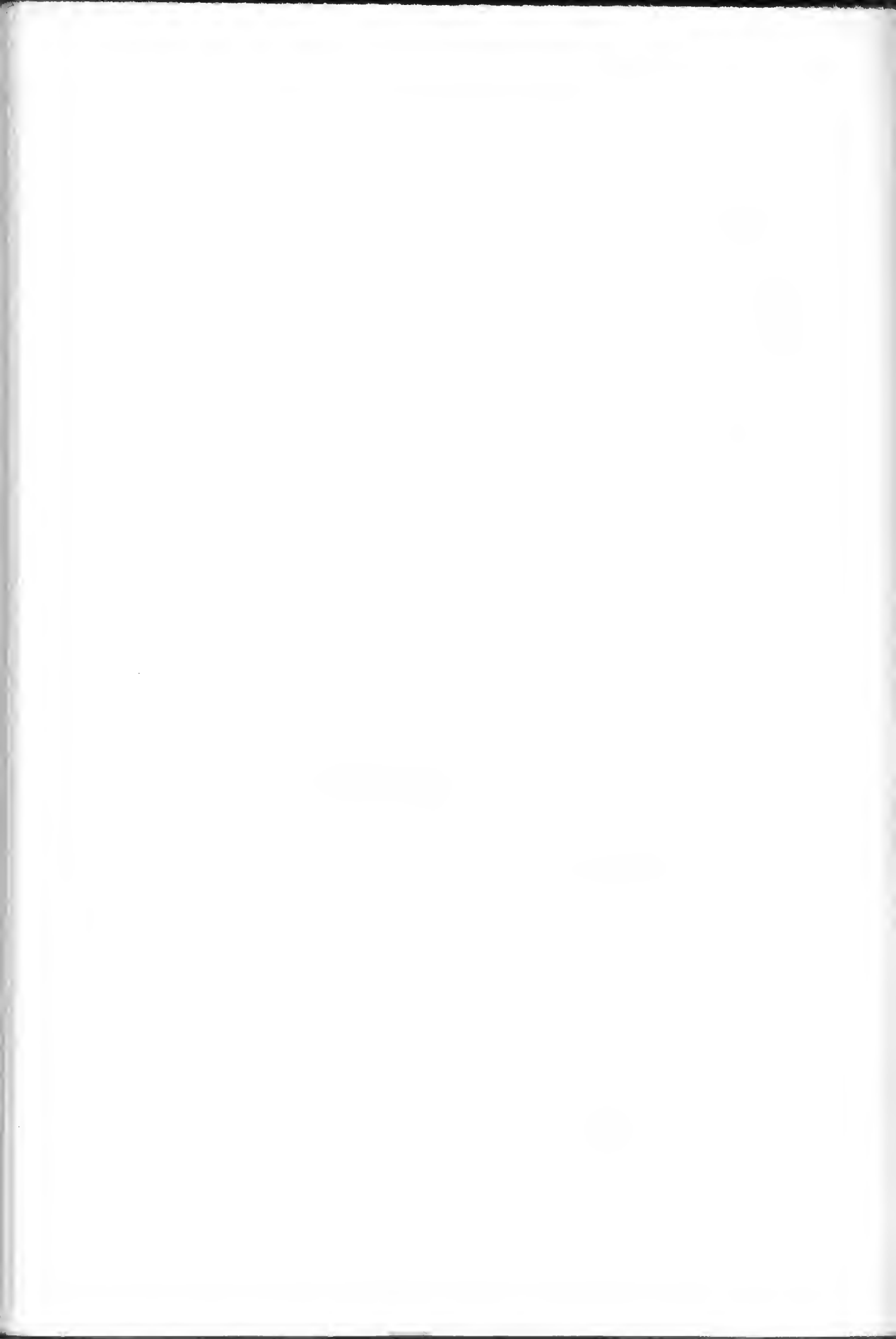
n	2	3	4	5	10	20	25	50	100	200	250	500	1000	2000	2500	5000	10.000
somma esp.	1	2	3	4	5	10	20	25	50	100	200	250	500	1000	2000	2500	5.000
	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
p	1	1	1	1	5	10	5	25	50	100	50	250	500	1000	500	2500	5.000
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

Come si vede la potenza 2000, ad esempio, può essere ottenuta con sole 14 operazioni.

BIBLIOGRAFIA

- [1] MIL-STD-721 - Definition of effectiveness terms for reliability, maintainability, human factors and safety, con rimando a MIL STD 882 B.
- [2] PAPOULIS, *Probability random variables and stochastic processes*. McGraw-Hill, Kogakusha, 1965.
- [3] GALEOTTI P., *Calcolo delle probabilità e statistica*. Ed. Levrotto & Bella, 1976.
- [4] LIEVENS G., *Securité des systemes*, Ecole National Supérieure de l'Aéronautique et de l'Espace, Sup. aéro, France.
- [5] FIORINI V., *Problemi di aeronavigabilità degli impianti*, Atti del Seminario sui sistemi aeronautici e spaziali, Torino 5-6 Giugno 1984, Ed. Levrotto & Bella, Torino.
- [6] ANTONA E., *Mathematical aspects of the probability evaluation of structural safety and NDI capability*.
- [7] ANTONA E., *Sicurezza, durata e affidabilità delle strutture aerospaziali*, Notizario AMMA, Torino, Maggio 1982.
- [8] ANTONA E., *Probability Principle and Static Load Safety*, Atti della Accademia delle Scienze di Torino, (to be published, presented December 1987).
- [9] ANTONA E., *Sicurezza e affidabilità nei veicoli aerei e spaziali*, Atti dell'Istituto di Navigazione, Roma 1975.

- [10] ANTONA E., *Safety in astronautic activities*, Atti della Accademia delle Scienze di Torino, Suppl. al vol. 122 (1988).
- [11] MIL-HDBK-338, APPENDIXA, *Dynamic Programming Approach to Reliability Allocation*, 15 october 1984.
- [12] CHIESA Sergio, *Affidabilità, sicurezza e manutenzione nel progetto dei sistemi*, Ed. CLUT, Torino.
- [13] VON MISES Richard, *Mathematical theory of probability and statistics*, Ed. Hilda Geiringer, Academic Press 1964, New York, London.
- [14] FRAGOLA Joseph, *Nasa efforts in risk and reliability data base development*, in «Research papers related to reliability, maintainability and risk assessment», Science applications international corporation, S.A.I.C., New York.
- [15] FRAGOLA Joseph R., SHEN Yu, *A Markov approach for addressing multiple failure probabilities in systems with repair*, in «Research papers related to reliability, maintainability and risk assessment», Science applications international corporation, S.A.I.C., New York and «17th inter-ram conference for the electric power industry».
- [16] FRAGOLA Joseph R., *Establishing a data base for an aerospace systems maintainability assessment*, in «Research papers related to reliability, maintainability and risk assessment», Science applications international corporation, S.A.I.C., New York.
- [17] PERKINS R. David, FRAGOLA Joseph R., *Propulsion reliability an historical perspective*, in «Research papers related to reliability, maintainability and risk assessment», Science applications international corporation, S.A.I.C., New York.
- [18] FRAGOLA Joseph R., *Reliability impact on earth observing satellites*, in «Research papers related to reliability, maintainability and risk assessment», Science applications international corporation, S.A.I.C., New York.
- [19] DOUGHERTY Ed M. Jr., FRAGOLA Joseph R., *Analogies in human performance between nuclear plant accident management and military combat*, in «Research papers related to reliability, maintainability and risk assessment», Science applications international corporation, S.A.I.C., New York.
- [20] FRAGOLA Joseph R., WOOD Stuart J., *Effectiveness evaluation of orbital observatories*, in «Research papers related to reliability, maintainability and risk assessment», Science applications international corporation, S.A.I.C., New York and «Proceedings 1975 annual reliability and maintainability symposium».
- [21] DIAMOND E.H. and FRAGOLA J.R., *Satellite long-life assurance, the impact of the Shuttle Era*, in «Research papers related to reliability, maintainability and risk assessment», Science applications international corporation, S.A.I.C., New York.
- [22] FRAGOLA J.R. and SPAHN J.F., *Economic models for satellite system effectiveness*, in «Research papers related to reliability, maintainability and risk assessment», Science applications international corporation, S.A.I.C., New York.



L'instabilità di gusci in materiale composito soggetti a compressione e torsione

Vittorio GIAVOTTO*, Carlo POGGI**,
Marios CHRYSSANTHOPOULOS***

***Sommario.** Viene presentato un progetto di ricerca sull'instabilità di gusci in materiale composito che prevede lo sviluppo di indagini sperimentali, teoriche e numeriche. Il confronto tra i risultati numerici e le prove sperimentali permetterà la validazione dei metodi che verranno successivamente utilizzati in un ampio studio parametrico. La ricerca si propone di produrre criteri per la progettazione di elementi strutturali a guscio in materiale composito soggetti a diverse combinazioni di carico. In questo lavoro vengono presentati i risultati delle prime prove sperimentali effettuate su cilindri soggetti al solo carico assiale. Viene inoltre brevemente riassunto il metodo utilizzato per valutare l'influenza delle diverse laminazioni sul comportamento critico e post-critico di gusci cilindrici compressi assialmente ed imperfetti.*

1.0. Introduzione

L'uso dei materiali compositi nelle strutture a guscio si è sviluppato principalmente nel campo aeronautico ed aerospaziale ma la mancanza di criteri progettuali di applicabilità generale per le strutture a guscio in materiale composito ha ostacolato un uso efficiente di tali materiali.

L'utilizzo di simulazioni numeriche è oggi comunemente accettato per l'analisi di strutture complesse come strumento alternativo alle più costose indagini sperimentali ma alcuni problemi strutturali, quali quelli relativi alla stabilità di strutture a guscio, possono essere risolti solo mediante una combinazione di indagini sperimentali, teoriche e numeriche.

La ricerca qui presentata si propone di migliorare le conoscenze del comportamento dei materiali compositi nelle strutture a guscio, di fornire risultati scientifici che permettano un migliore sfruttamento delle

* Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano.

** Dipartimento di Ingegneria Strutturale, Politecnico di Milano.

*** Department of Civil Engineering Imperial College - London.

proprietà di tali materiali valutando l'influenza dei processi produttivi sulle funzionalità della struttura. I risultati numerici e sperimentali prodotti durante la ricerca formeranno una base di dati utili allo sviluppo di normative Europee per la progettazione di strutture a parete sottile in materiale composito soggette a diverse combinazioni di carico.

La ricerca si sta sviluppando secondo le seguenti fasi:

- valutazione delle proprietà statistiche delle imperfezioni geometriche registrate su alcuni gruppi di campioni cilindrici costruiti con materiale composito e con diverse sequenze di laminazione;
- sviluppo di «modelli caratteristici» di imperfezioni che saranno utilizzati nei programmi di calcolo per la definizione delle capacità portante di ogni elemento;
- prove sperimentali su alcuni gruppi di provini cilindrici con differenti sequenze di laminazione e combinazioni di carico;
- registrazione durante le prove sperimentali del comportamento pre e post-critico e del meccanismo di collasso per un successivo confronto con gli analoghi risultati numerici;
- definizione dell'affidabilità delle prove sperimentali di instabilità mediante il confronto dei risultati ottenuti su campioni nominalmente identici;
- confronto tra i valori di collasso previsti mediante i metodi analitici e numerici ed i risultati delle prove sperimentali al fine di validare i modelli numerici utilizzabili in un successivo studio parametrico;
- sviluppo, sulla base dei risultati numerici e sperimentali, di criteri progettuali per cilindri in materiale composito soggetti ad azione assiale, torsione e ad una loro combinazione di carico.

2.0. Programma sperimentale

2.1. Geometria dei campioni cilindrici e proprietà del materiale

I campioni cilindrici utilizzati in questo lavoro sono costituiti di lamine di tessuto di Kevlar impregnato in una matrice di resina epossidica. I modelli cilindrici (Fig. 1) hanno raggio pari a 350 mm e lunghezza totale di 700 mm. Si noti che i cilindri presentano due parti di maggiore spessore nella parte superiore ed inferiore per facilitare il fissaggio del campione nell'apparecchiatura di prova (Fig. 2). Di conseguenza la lunghezza reale del cilindro in parete sottile risulta di 550 mm.

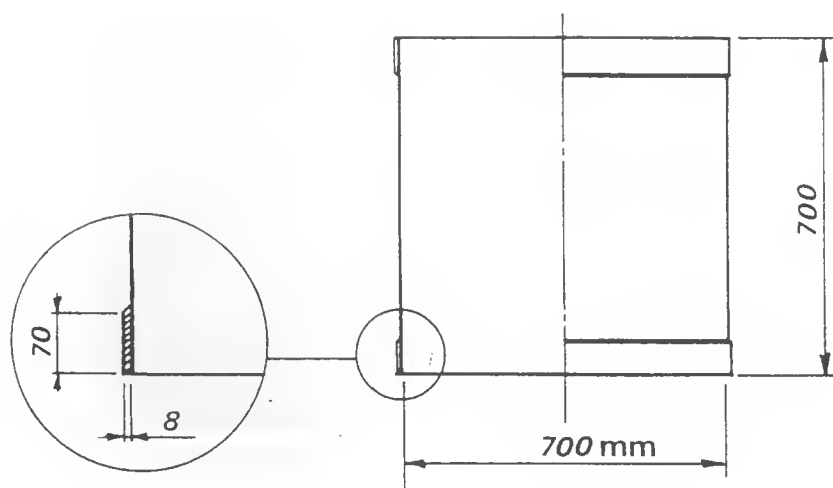


Figura 1 - Geometria dei campioni cilindrici.

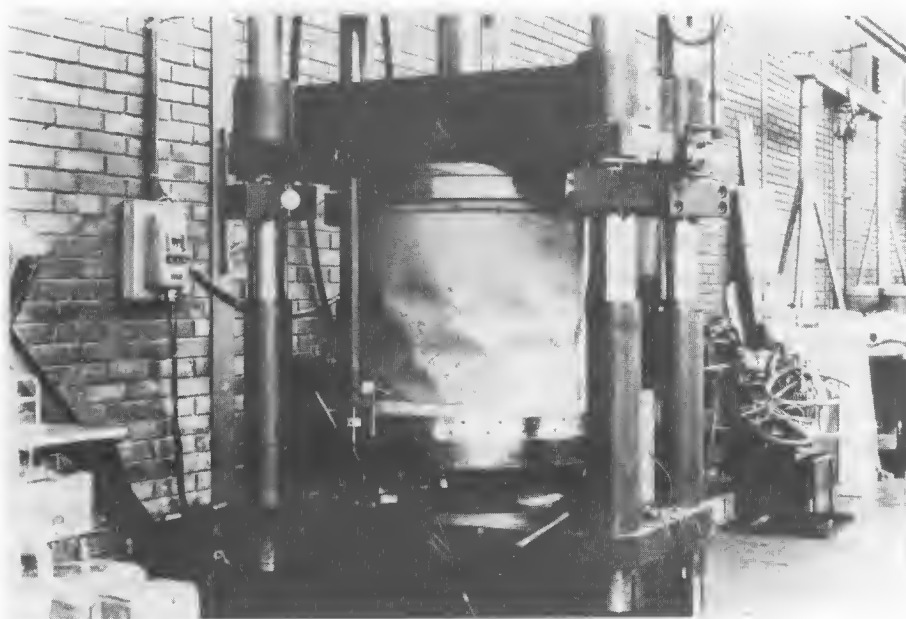


Figura 2 - Apparecchiatura di prova.

Le caratteristiche elastiche di ogni lamina sono le seguenti

$$E_1 = E_2 = 23450 \text{ MPa}$$

$$\nu_{12} = 0.2$$

$$G_{12} = G_{23} = G_{31} = 1520 \text{ MPa}$$

gli indici 1 e 2 si riferiscono agli assi x_1 ed x_2 diretti come le fibre di ogni singola lamina mentre l'asse x_3 è normale al piano medio di ogni lamina. Lo spessore di ogni lamina è di 0.26 mm.

Attualmente sono disponibili campioni cilindrici costruiti con diverse laminazioni:

- A) 16 cilindri con laminazione cross-ply con 4 lamine disposte a $(0^\circ/90^\circ)_s$ rispetto all'asse del cilindro, spessore totale = 1.04 mm;
- B) 14 cilindri con laminazione angle-ply con lamine disposte a $(\pm 45)_s$ rispetto all'asse del cilindro, spessore totale = 1.04 mm;
- C) 3 gruppi di 4 cilindri «quasi-isotropi» costituiti da 8 lamine con le seguenti sequenze di laminazione $(45/-45/0/90)_s$, $(45/0/90/45)_s$, $(0/90/45/-45)_s$ (spessore totale = 2.08 mm).

Essendo ogni lamina ortotropa è chiaro che i cilindri con laminazione «cross-ply» ed «angle-ply», essendo costituiti da più lamine con le stesse caratteristiche, presentano le stesse caratteristiche elastiche di una lamina con spessore quadruplo.

2.2. *L'apparecchiatura di prova sperimentale*

L'apparecchiatura di prova utilizzata per le sperimentazioni sui campioni è rappresentata in fig. 2. L'azione assiale è fornita da una pompa idraulica ma il reale carico assiale agente sul campione sottoposto a prova è regolato da quattro fermi a vite agenti negli spigoli della testa superiore della macchina di prova. È così possibile regolare il processo di carico in controllo di spostamento con un'ottima accuratezza. Il campione cilindrico da provare è vincolato inferiormente ad un piatto metallico supportato da tre rulli che, nel caso venga applicata solo azione assiale, sono appoggiati su un piano orizzontale (Fig. 3). Nel caso di combinazioni di azione assiale e torsione i suddetti rulli possono ruotare su superfici inclinate. La torsione pura è applicata al provino mediante un diverso sistema indipendente di carico. Durante le prove vengono registrati gli spostamenti ed il carico assiale in corrispondenza dei tre supporti per verificare l'uniformità degli spostamenti e la centralità del carico.

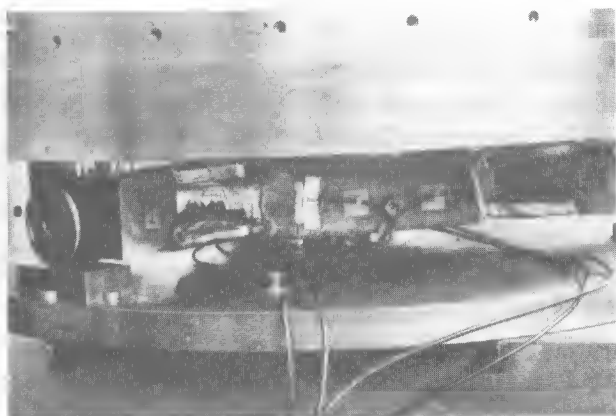


Figura 3 - Supporti a rullo per il caso di carico di pura torsione e compressione.

2.3. *La mappatura delle imperfezioni geometriche*

La registrazione delle imperfezioni di forma presenti sui provini cilindrici è stata effettuata mediante un'attrezzatura appositamente studiata. Gli scostamenti delle superfici interna ed esterna del cilindro da una superficie ideale cilindrica sono stati registrati mediante due trasduttori. L'acquisizione dei dati è regolata da un computer ed i dati vengono memorizzati per successive elaborazioni numeriche. L'apparecchiatura permette la registrazione delle imperfezioni nelle due direzioni assiale e circonferenziale secondo una scansione desiderata. In particolare, nel caso in esame si è proceduto alla registrazione delle imperfezioni su una rete di punti con intervallo assiale pari ad 1 cm. e circonferenziale pari a 2 cm. Tale operazione si riconduce quindi a 46 letture assiali per ogni meridiano e 110 letture per ogni circonferenza per la superficie esterna e ad altrettante per la superficie interna del cilindro per un totale di 10120 registrazioni per ogni cilindro.

3.0. *Analisi delle imperfezioni geometriche*

3.1. *Analisi biarmonica delle imperfezioni geometriche*

La sensibilità alle imperfezioni è a lungo stata riconosciuta come la causa principale delle discrepanze riscontrate tra i valori sperimentali ed

valori teorici dei carichi di collasso di strutture a guscio in materiale isotropo. Questo fenomeno è particolarmente accentuato nel caso di gusci cilindrici soggetti a compressione. Negli ultimi anni sono stati effettuati numerosi studi relativi all'influenza delle imperfezioni effettuando accurate misurazioni delle imperfezioni presenti su provini cilindrici ed anche su campioni in scala reale da sottoporre a prove sperimentali [2]. Nella maggior parte di questi studi è stato adottato un metodo standard di analisi dei dati basato sul concetto di «best-fit cylinder» [3]. I valori registrati direttamente sul provino, ovvero ottenuti dalle letture dei trasduttori, formano i dati di ingresso di una procedura che valuta le dimensioni del cilindro che meglio interpola i valori registrati sulla griglia precedentemente definita ed, in una seconda fase, ricalcola i valori delle imperfezioni come scostamenti da questa superficie fittizia perfetta. Questa procedura, che esclude dalle letture eventuali disallineamenti o rotazioni rigide del campione, ha permesso di costruire una base di dati unificata per le imperfezioni delle strutture a guscio e si è mostrata particolarmente utile nel caso di studi comparativi tra serie di campioni analoghi.

Dopo aver operato la suddetta procedura di «best-fit» è possibile analizzare le caratteristiche delle imperfezioni risultanti mediante un'analisi armonica bidimensionale e calcolare i valori dei coefficienti di Fourier di una serie del tipo

$$(1) \quad w_0(x, \theta) = \sum_{m=1}^{\bar{m}} \sum_{n=0}^{\bar{n}} \xi_{mn} \sin \frac{m\pi x}{L} \sin (n\theta + \phi_{mn})$$

dove ξ_{mn} e ϕ_{mn} sono coefficienti di Fourier ottenuti dalla misura in punti discreti di una funzione imperfezione, $w_0(x, \theta)$, definita sulla parete del cilindro, ovvero in un campo ($0 \leq x \leq L$ $0 \leq \theta \leq 2\pi$).

Si noti che l'espressione sopra riportata rappresenta uno sviluppo in serie di semionde di seno in direzione assiale implicando che il valore dell'imperfezione è nullo ai due estremi. Sebbene ciò non sia corretto l'errore introdotto è limitato alle regioni terminali del campo in questione e non è significativo, a condizione che il numero dei termini corrispondenti dello sviluppo in serie non sia troppo piccolo. A verifica di ciò sono stati confrontati i risultati ottenuti mediante una serie di semionde di seno o di coseno utilizzando la seguente funzione di errore

$$(2) \quad e = \frac{1}{N} \sum_{i=1}^N (w_i^{BF} - w_i^F)^2$$

dove w_i^{BF} è il valore dell'imperfezione nel punto i dopo la procedura di 'best-fit', w_i^F è il valore dell'imperfezione nel punto i utilizzando la serie di Fourier ed N è il numero totale di punti sul cilindro in cui sono state effettuate le letture delle imperfezioni. Anche l'utilizzo di una serie completa di seni e coseni ha confermato la bontà e la convenienza della scelta effettuata. È stato valutato inoltre l'errore in corrispondenza dei punti relativi alle massime imperfezioni verso l'interno e verso l'esterno del cilindro. In generale la serie comprendente semionde di seno in direzione assiale è stata la più conveniente in termini di efficienza e precisione.

L'utilizzo di tale procedura per l'analisi delle imperfezioni geometriche permette la valutazione dell'ampiezza di singoli modi di imperfezione ed il loro utilizzo in strumenti analitici o numerici per uno studio parametrico del carico limite di strutture a guscio. Questo metodo si rivela particolarmente vantaggioso solo nel caso in cui siano disponibili le registrazioni delle imperfezioni effettuate su gruppi di campioni costruiti con tecniche simili. È allora possibile applicare tecniche statistiche sui valori calcolati dei coefficienti per arrivare ad un modello di «imperfezione caratteristica» associata alla particolare tecnica costruttiva utilizzata [4]. I risultati di tale analisi sono utili per quantificare il carico di collasso di cilindri in presenza di modelli di imperfezioni mono e multimodali [5]. L'uso di metodi probabilistici per il calcolo dell'affidabilità di strutture a guscio con imperfezioni random è stato studiato in modo approfondito in [6] da Elishakoff e Arbocz.

Il presente programma di ricerca ben si adatta a questo tipo di analisi in quanto tra i campioni da sottoporre a prove si distinguono due serie sufficientemente ampie di cilindri «nominalmente identici». In particolare la prima di tali serie è costituita da sedici provini «cross-ply» $(0^\circ/90^\circ)_s$ e la seconda comprende quattordici campioni «angle-ply» $(+45^\circ/-45^\circ)_s$. Si ritiene che il confronto dei modelli caratteristici di imperfezione derivati da diverse sequenze di laminazione possa essere sufficiente per la definizione di tolleranze geometriche specifiche e dei relativi metodi di ispezione su strutture reali [7].

3.2. *Analisi statistica delle imperfezioni geometriche*

L'analisi di «best-fit» sopra descritta è stata applicata a tutte le registrazioni di imperfezioni relative ai cilindri in esame. In una successiva fase è stata sviluppata un'analisi di Fourier utilizzando le funzioni (1). Il modello delle imperfezioni di ogni cilindro è risultato quindi descritto

da un gruppo di coefficienti ξ_{mn} e ϕ_{mn} , con $\bar{m}=20$ ed $\bar{n}=40$. Questi coefficienti sono stati successivamente oggetto di un'indagine statistica al fine di identificare caratteristiche comuni utili alla definizione di modelli caratteristici di imperfezione.

Le Fig. 4a, b mostrano le superfici tipiche dei cilindri imperfetti (dopo la procedura di «best-fit») ottenuti per due cilindri della serie A (cross-ply) e B (angle-ply). È interessante notare che le due superfici interne mostrano caratteristiche comuni mentre le superfici esterne sono fortemente influenzate dall'orientamento delle fibre. I picchi rilevabili sulle superfici esterne sono infatti il risultato di variazioni locali di spessore dovuti alla sovrapposizione di lamine e rispecchiano le diverse laminazioni delle due serie. Per questo motivo tali picchi non sono analizzabili come singole imperfezioni ma saranno comunque oggetto di successivi studi per esaminarne l'eventuale influenza sul carico limite.

Il confronto tra le superfici interne ed esterne, espresso in termini di coefficienti di Fourier, è riportato nelle Fig. 5a, b per la serie A di cilindri (cross-ply). Come si può notare le curve rappresentative dei valori medi hanno lo stesso andamento fino a valori di n circa uguali a 15 mentre divergono per valori superiori. Ciò dimostra che i modi di imperfezione corrispondenti a onde circonferenziali corte sono presenti solo sulla superficie esterna a causa delle variazioni localizzate di spessore. Analoghe conclusioni, qui non riportate, sono state dedotte per la serie B di cilindri. Si può quindi concludere che per l'analisi degli effetti delle imperfezioni iniziali è sufficiente effettuare misure sulla superficie interna.

Tornando ai risultati relativi alla superficie interna riportati in Fig. 5b, è evidente come i modi dominanti siano associati a onde circonferenziali lunghe e come gli estremi dei valori per ogni numero di onde circonferenziali n abbiano un andamento in accordo con i relativi valori medi. I cilindri della serie B (angle-ply) presentano caratteristiche simili. Questo implica che la disposizione delle lamine, in cilindri costruiti con lo stesso processo, non ha influenza significativa sulle caratteristiche delle imperfezioni. È inoltre interessante notare che la predominanza di modi con onda circonferenziale lunga è stata osservata anche in cilindri isotropi [4].

L'ampiezza media delle singole imperfezioni può essere modellata utilizzando un'espressione della forma

$$E(\xi_{mn}) = e^{\alpha n^{\beta}}$$

dove α e β sono costanti valutate sulla base dei valori medi dei campioni.

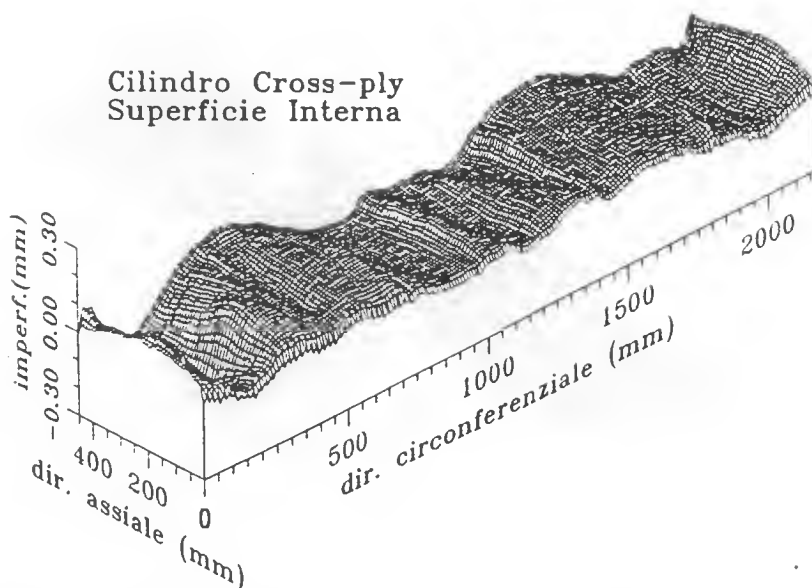
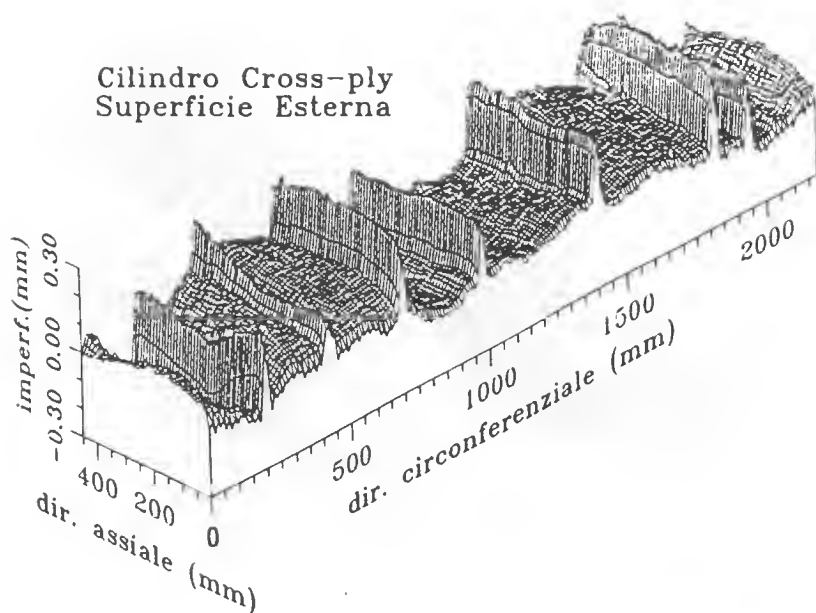
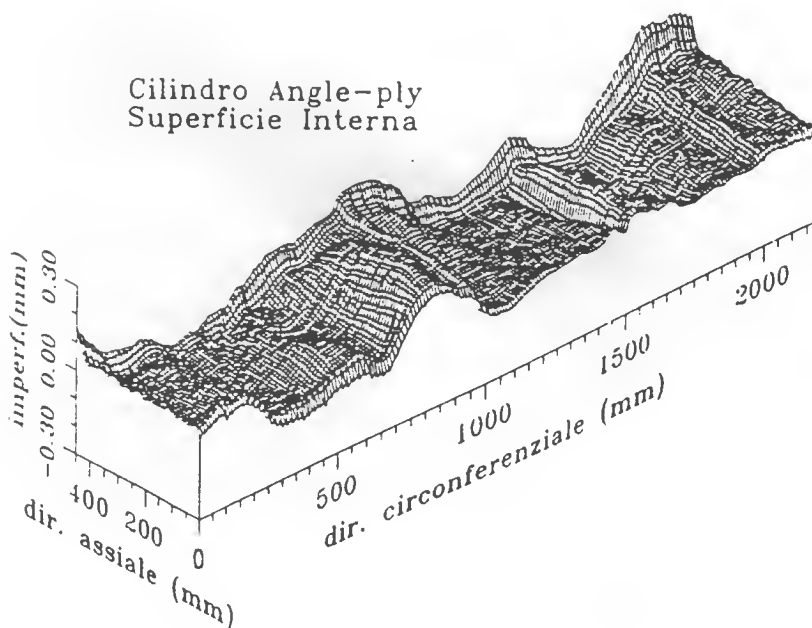
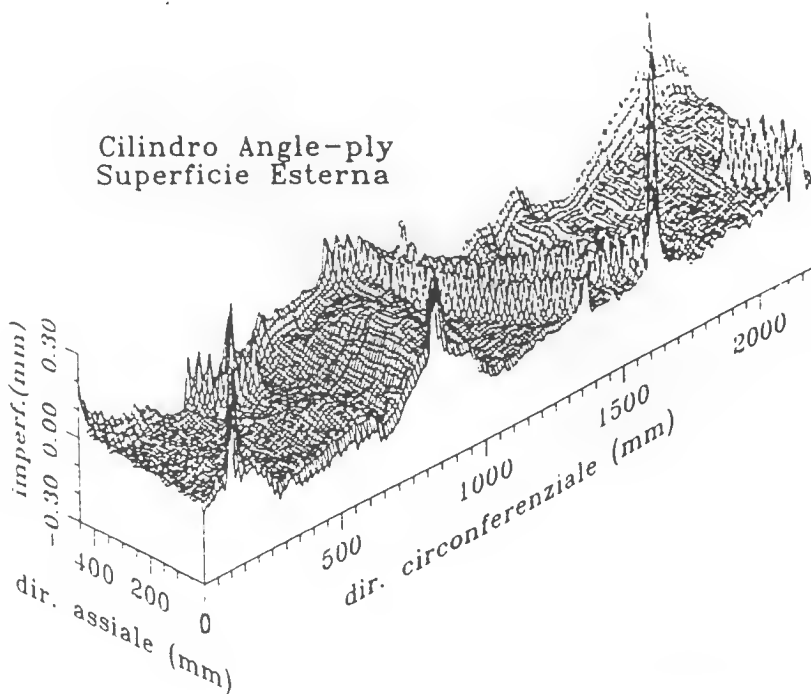


Figura 4 - Tipica superficie rappresentante le imperfezioni di un cilindro
a) cross-ply



b) angle-ply.

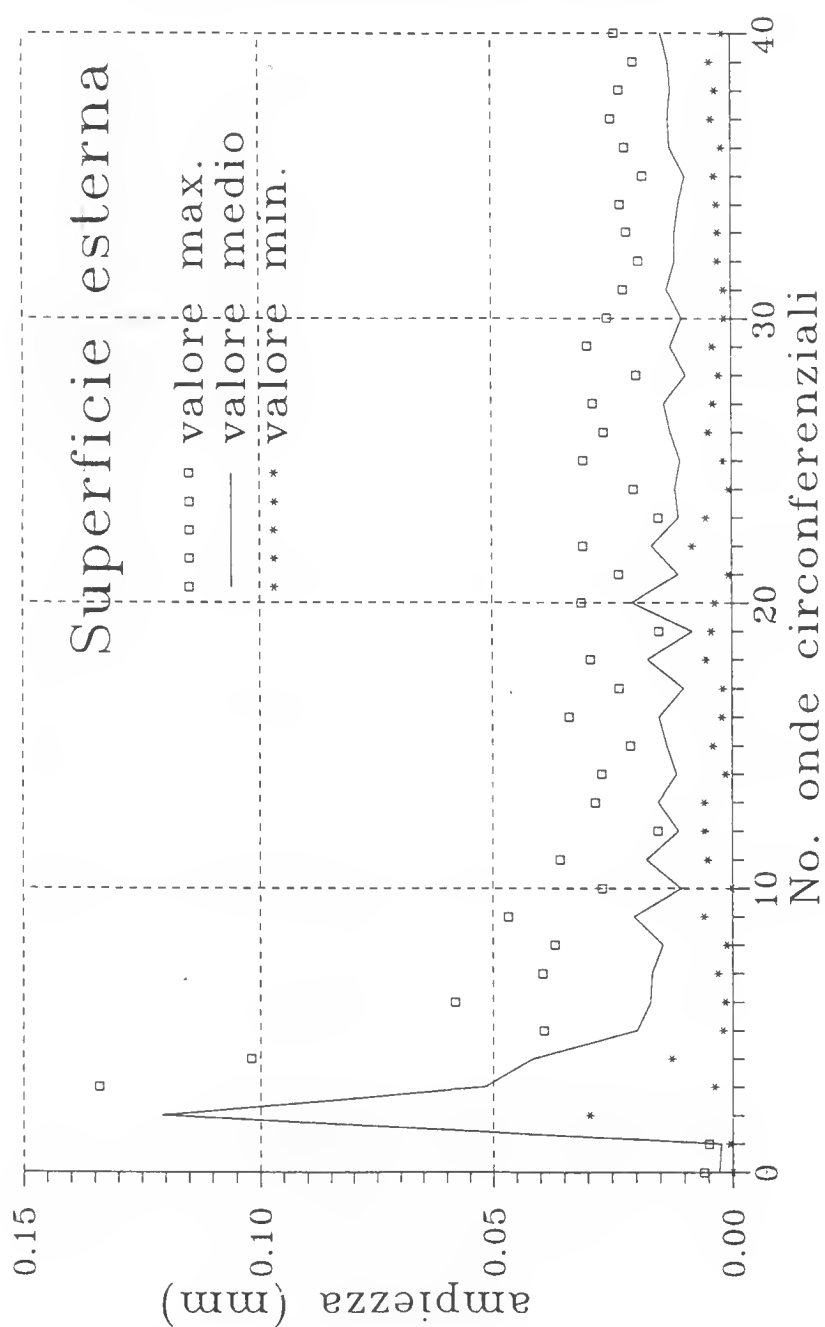


Figura 5 - Analisi dei valori medi delle ampiezze dei modi di imperfezione per
a) superficie esterna del cilindro (serie A, $m=1$)

La Fig. 6 presenta i risultati di un'analisi di variabilità per cilindri cross-ply. Si nota che i coefficienti di variazione (deviazione standard/media) non presentano una variabilità significativa. Un'espressione che lega l'ampiezza media di ogni modo alla sua deviazione standard può essere ottenuta per mezzo di una regressione per esempio nel seguente modo

$$(4) \quad \sigma(\xi_{mn}) = \gamma E(\xi_{mn})$$

dove γ è una costante.

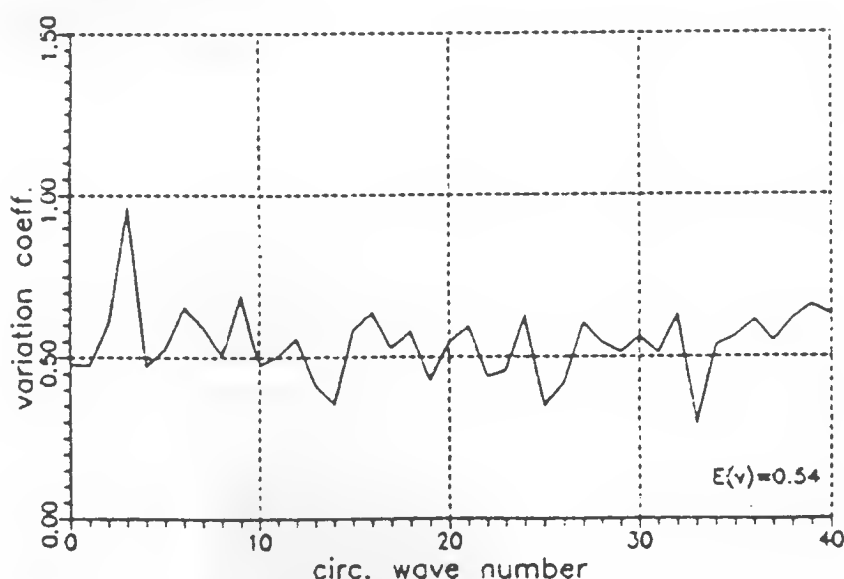


Figura 6 - Coefficienti di variazione dell'ampiezza dei modi di imperfezione (serie A, $m=1$).

Arbocz ha dimostrato che lo sviluppo di tali semplici espressioni che contengono le caratteristiche salienti delle ampiezze delle imperfezioni permette un confronto immediato dei modelli caratteristici delle imperfezioni generate da diverse tecniche costruttive.

Oltre all'analisi statistica sopra riportata è stata sviluppata anche un'analisi di correlazione tra le ampiezze modali. I coefficienti di correlazione sono risultati alti nel caso di modi con lo stesso numero di onde circolari ed un numero dispari di onde assiali ovvero $\rho(\xi_{1n}, \xi_{mn})$ per $m=3,5,\dots$. La correlazione tra i modi con lo stesso numero di onde longitudinali è invece risultata molto più bassa. Infine, per ottenere un

completo modello probabilistico della variabilità delle ampiezze modali, si è verificato quale tra le più comuni distribuzioni di probabilità meglio approssimava i risultati ottenuti. Sono stati determinati a tale scopo i coefficienti delle distribuzioni Weibull e log-normale per le ampiezze modali dominanti.

I risultati dell'analisi statistica hanno permesso di concludere che è possibile ritrovare andamenti comuni nelle distribuzioni delle imperfezioni nelle due serie di campioni esaminati. È stato quindi possibile sviluppare vari modelli che permettono di descrivere superfici caratteristiche delle imperfezioni che saranno utilizzati in successivi studi parametrici numerici ed analitici per giungere alla definizione di criteri progettuali per l'instabilità di cilindri in materiale composito soggetti a compressione e momento torcente.

4.0. Risultati delle prove sperimentali

I campioni cilindrici disponibili sono destinati ad essere sottoposti a prove di compressione, di torsione ed a combinazioni di questi [9]. Fino ad ora sono stati provati a compressione semplice quattro cilindri appartenenti al gruppo A (cross-ply). In fig. 7 è riportato un tipico diagramma carico-spostamento medio in direzione assiale ottenuto su uno stesso provino in seguito a due prove successive. È immediato notare come le caratteristiche elastiche del materiale non siano alterate da precedenti processi di carico assicurando la quasi coincidenza dei due percorsi post-critici. Si nota inoltre che il valore del carico sopportabile nel ramo post critico risulta pari a circa il 60% del carico limite raggiunto. I risultati di prove effettuate su campioni analoghi sono riportati in Tabella 1 a confronto con il valore teorico del carico critico ottenibile per un cilindro perfetto. Tutti i risultati sperimentali ottenuti sono inclusi nell'intervallo tra 27 e 29 KN. Inoltre i diagrammi carico-spostamento ottenuti per gli altri cilindri, analogamente, presentano un ramo pre-critico decisamente lineare ed un percorso post-critico lentamente decrescente simile a quello rappresentato in Fig. 7. Dall'esame dei valori riportati in Tabella 1 si può notare che l'effetto della presenza di imperfezioni geometriche in cilindri cross-ply non è particolarmente marcato in quanto il coefficiente riduttivo ad esso corrispondente si aggira intorno al valore 0.70. In Fig. 8 è riportata la deformata critica di uno dei provini sottoposti a prova di compressione.

I risultati del programma sperimentale in corso permetteranno la definizione di curve di sensibilità alle imperfezioni per ogni diversa se-

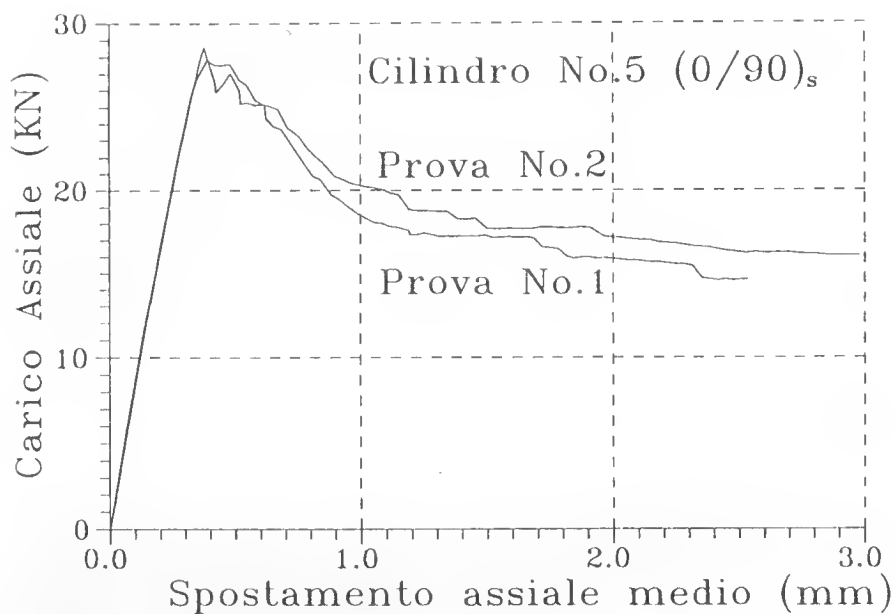


Figura 7 - Curve carico-spostamento assiale di una prova sperimentale (cilindro cross-ply).



Figura 8 - Deformata tipica di un cilindro cross-ply soggetto a compressione.

quenza di laminazione di questi cilindri nominalmente identici [10]. Contemporaneamente alle prove sperimentali ogni singolo cilindro, soggetto alle imperfezioni di forma precedentemente registrate, è analizzato mediante metodi numerici. Si prevede inoltre di studiare in dettaglio, sempre con modelli numerici, l'effetto della variazione di spessore dovuta alla sovrapposizione delle fibre delle varie lamine costituenti il cilindro. Lo scopo è quello di giungere, sulla base dei risultati sperimentali integrati da quelli numerici, ad un'espressione generale che definisca la sensibilità alle imperfezioni di cilindri in materiale composito soggetti a compressione.

5.0. Analisi teorica e numerica

L'analisi numerica è sviluppata mediante due differenti e complementari strumenti di calcolo: un programma ad elementi finiti per l'analisi di strutture a guscio in materiale composito ed una procedura di calcolo basata sull'analisi perturbativa delle equazioni governanti il problema.

Quest'ultima tecnica risolutiva permette di giungere ad espressioni esatte del carico critico e di un coefficiente governante il ramo post-critico di un cilindro con una qualsiasi disposizione delle fibre. È possibile inoltre un esame accurato dell'influenza dei singoli modi di imperfezione, siano essi assialsimmetrici o asimmetrici, e descritti da una qualsiasi funzione analitica, che può essere una delle funzioni trigonometriche utilizzate per l'analisi biarmonica delle imperfezioni descritta nel paragrafo 3.1. Tale approccio è però evidentemente applicabile alle sole strutture con geometria assialsimmetrica.

L'uso degli elementi finiti permette invece una simulazione delle prove sperimentali includendo una dettagliata descrizione delle imperfezioni geometriche registrate. Il confronto tra prove numeriche e sperimentali è sviluppato in termini sia di curve carico spostamento che di sforzi in punti significativi della struttura. Lo scopo principale di questi confronti è evidentemente quello di verificare la capacità del modello ad elementi finiti di simulare una prova sperimentale e di produrre il valore del carico limite di un guscio cilindrico in materiale composito. L'esito positivo di tale confronto permetterà una successiva indagine parametrica per tenere in conto più ampi campi di variazione dei parametri geometrici e delle condizioni di vincolo.

L'indagine numerica permetterà quindi di integrare i risultati sperimentali e di formulare espressioni analitiche per la valutazione del carico limite che siano idonee alla pratica progettuale. Qui nel seguito verrà

brevemente descritta la tecnica perturbativa utilizzata per uno studio teorico dei gusci cilindrici in materiale composito e richiamati alcuni risultati ottenuti mediante un programma ad elementi finiti descritto in [9].

5.1. *Il metodo perturbativo*

Il problema della stabilità dei gusci in composito è stato affrontato mediante l'applicazione della teoria generale della stabilità mediante la quale è possibile ottenere una soluzione in forma asintotica del comportamento nonlineare dei gusci soggetti ad imperfezioni geometriche [12]. L'applicazione di tale metodo e la soluzione delle equazioni differenziali risultanti implica però la manipolazione di espressioni analitiche lunghe e complesse [11]. Tale compito può essere facilitato dall'utilizzo di programmi per la manipolazione simbolica di espressioni matematiche che permettono di ottenere con relativa facilità soluzioni esatte delle equazioni in oggetto, senza rischi di introdurre errori. In particolare per questo lavoro è stato utilizzato un programma per l'analisi simbolica utilizzabile anche su calcolatori di medie dimensioni e capacità.

Le equazioni che governano il comportamento di gusci cilindrici in materiale composito sono le seguenti

$$\begin{aligned} L_1(\phi) - L_2(w) &= -1/2 N_4(w, w) \\ (5a, b) \quad L_3(w) + L_2(\phi) &= N_4(w, \phi) \end{aligned}$$

dove w è la componente di spostamento normale alla superficie del guscio (direzione z) (Fig. 9), ϕ è una funzione potenziale degli sforzi membranali mentre L_1 , L_2 , L_3 e N_4 sono operatori differenziali definiti in Appendice.

5.2. *Carico critico e comportamento post-critico iniziale*

L'applicazione di un metodo perturbativo alle equazioni (5a, b) permette di individuare i punti biforcativi lungo il percorso di equilibrio ed avere informazioni locali sui percorsi diramati. Una soluzione del problema può essere infatti espressa in forma asintotica

$$\begin{aligned} \lambda &= \lambda_c + \lambda_1 \alpha + \lambda_2 \alpha^2 + \dots \\ (6a, \vartheta) \quad w &= w_0 + w_1 \alpha + w_2 \alpha^2 + \dots \\ \phi &= \phi_0 + \phi_1 \alpha + \phi_2 \alpha^2 + \dots \end{aligned}$$

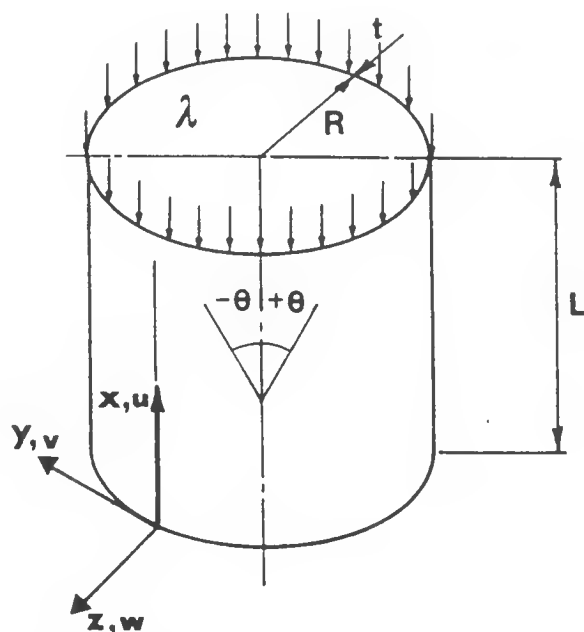


Figura 9 - Geometria del cilindro.

Sostituendo tali espressioni nelle (5a, b) e nelle relative condizioni al contorno [12] ed annullando identicamente i termini ad egual potenza di α (in questo caso limitati al secondo ordine) si ottengono i seguenti problemi lineari

$$(7a, b) \quad \begin{aligned} L_1(\phi_1) - L_2(w_1) &= 0 \\ L_3(w_1) + L_2(\phi_1) + \lambda_C w_{1,XX} &= 0 \end{aligned}$$

Tale problema si riduce ad un problema agli autovalori la cui soluzione può essere assunta del tipo

$$(8a, b) \quad \begin{aligned} w_1 &= \sin \frac{m\pi x}{L} \sin \frac{ny}{R} \\ \phi_1 &= C_{mn} \sin \frac{m\pi x}{L} \sin \frac{ny}{R} \end{aligned}$$

dove

$$C_{mn} = \frac{-\frac{m^2 \pi^2}{L^2 R} + \hat{B}_{mn}}{\hat{A}_{mn}}$$

mentre $\hat{A}_{mn}$ e $\hat{B}_{mn}$ sono riportati in Appendice.

Il valore del carico critico λ_C è dato dalla seguente espressione

$$(9) \quad \lambda_C = -\frac{C_{mn}}{R} + \left(\frac{L}{m\pi}\right)^2 [\hat{D}_{mn} + C_{mn} \hat{B}_{mn}]$$

$\hat{D}_{mn}$ è riportato in Appendice.

Le equazioni perturbative del secondo ordine risultano

$$(10a, b) \quad \begin{aligned} L_1(\phi_2) - L_2(w_2) &= -1/2 N_4(w_1, w_1) \\ L_3(w_2) + L_2(\phi_2) + \lambda w_{2,xx} &= N_4(w_1, \phi_1) \end{aligned}$$

Questo sistema ammette soluzioni del tipo

$$(11a, b) \quad \phi_2 = \sum_{i=1}^{\infty} \left\{ b_{0i} + b_{2i} \cos \frac{2ny}{R} \right\} \sin \frac{i\pi x}{L}$$

$$w_2 = \sum_{i=1}^{\infty} \left\{ c_{0i} + c_{2i} \cos \frac{2ny}{R} \right\} \sin \frac{i\pi x}{L}$$

Da cui è possibile ricavare il coefficiente del second'ordine

$$(12) \quad \lambda_2 = \sum_{i=1}^{\infty} \frac{4n^2}{\pi R^2} \left[\frac{2i(b_{0i} + 2c_{0i} C_{mn})}{(i^2 - 4m^2)} + \frac{b_{2i} + 2c_{2i} C_{mn}}{i} \right]$$

che permette di valutare il valore del carico limite e la sensibilità alle imperfezioni di una qualsiasi struttura cilindrica in materiale composito che presenti un modo critico singolo [12].

5.3. Risultati ed esempi

Vengono qui presentati alcuni dei risultati ottenuti utilizzando la teoria sopra esposta, relativi a cilindri soggetti al solo carico assiale. In Fig. 10 sono riportati i risultati relativi allo studio sulla variazione del valore del carico critico, eq. (9), in dipendenza dell'angolo di inclinazione delle fibre θ (Fig. 11). Si è assunto che i termini B_{ij}^* siano nulli e che i termini di accoppiamento siano trascurabili. Ciò è vero solo per valori di $\theta = 0^\circ$ e 45° e per laminati simmetrici. Si noti come il carico critico per valori dell'angolo θ pari a 0° e 45° siano molto vicini ma i corrispondenti modi critici siano rispettivamente asimmetrico (per laminazione cross-ply) ed assialsimmetrico (per laminazione angle-ply a 45°). Si nota inoltre come il modo critico asimmetrico sia il dominante per laminazioni fino a circa 23° e viceversa il modo critico diventi assialsimmetrico per laminazioni con angoli θ compresi tra 23° e 45° . È inoltre evidente

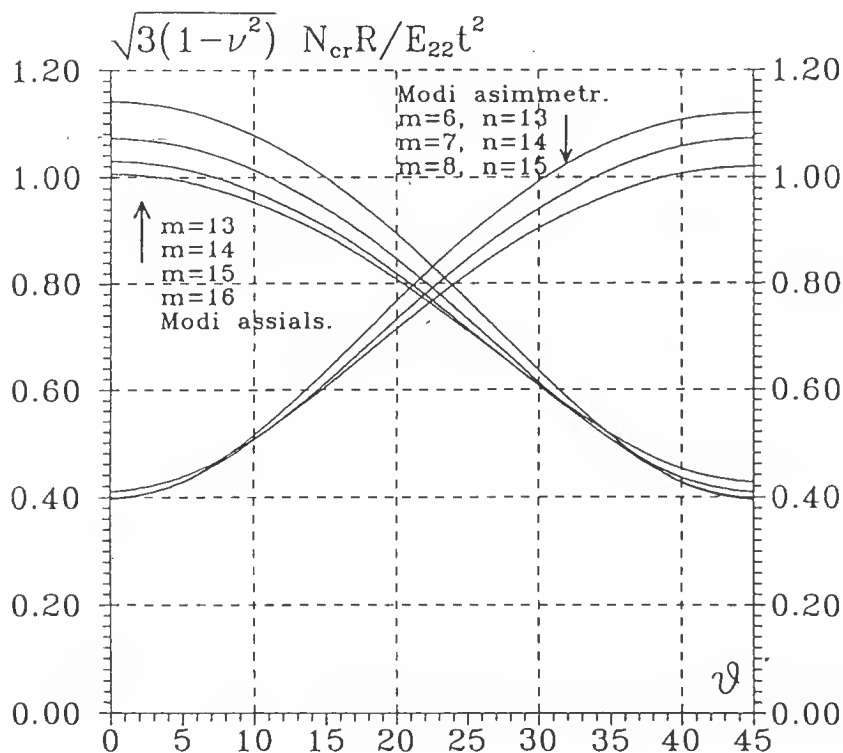


Figura 10 - Carichi critici per cilindri con diversi orientamenti delle fibre.

come il valore del carico critico più elevato corrisponda ad una laminazione $\theta = 23^\circ$ e tale configurazione sia identificabile come la più conveniente sotto tale aspetto. È comunque necessario sottolineare che in corrispondenza di tale valore sono presenti diversi modi critici quasi-simultanei che, come noto, producono una maggiore sensibilità alle imperfezioni.

In Fig. 11a, b sono riportate le mappe tridimensionali dei valori del carico critico adimensionalizzato, relativo a diversi modi critici, per laminazioni angle-ply ($\pm 45^\circ$)_s e cross-ply ($0^\circ/90^\circ$)_s. È evidente come in entrambi i casi sia ben individuabile un singolo modo critico. Contrariamente a ciò i cilindri quasi-isotropi con una sequenza di laminazione ($45^\circ/-45^\circ/0^\circ/90^\circ$)_s mostrano modi critici simultanei che rendono il comportamento critico di tale cilindro molto simile ad un cilindro isotropo (Fig. 11c, d). In tali figure è immediato riconoscere che i valori minimi del carico critico si distribuiscono secondo un cerchio simile al cosiddetto cerchio di Koiter. Questa caratteristica può essere molto importante nella determinazione della sensibilità alle imperfezioni di ciascuna diversa laminazione. Si prevede pertanto di esaminare in dettaglio il comportamento di cilindri quasi-isotropi che si ritengono molto sensibili alle imperfezioni analogamente ai cilindri isotropi. L'adozione di tale laminazione che notoriamente permette, a parità di spessore globale della superficie del cilindro, di ottenere un valore più elevato del carico critico, avrebbe lo svantaggio di presentare un andamento post-critico molto ripido.

L'analisi sarà in seguito estesa, seguendo un procedimento del tutto analogo, al caso di sollecitazione di torsione o ad azione combinata di compressione e torsione.

5.4. *Il modello ad elementi finiti*

Le sopracitate analisi sono state condotte sotto l'ipotesi di modo critico singolo. Come noto nei casi reali si riscontra invece un comportamento corrispondente alla combinazione di più modi singoli. Tale fenomeno necessita di essere studiato mediante una analisi multimodale o ad elementi finiti. Qui di seguito sono commentati i risultati mediante un'analisi ad elementi finiti eseguita mediante il programma descritto in [9]. I risultati numerici sono riportati in Tabella 2 a confronto con quelli ottenuti dall'eq. (9). Le deformate critiche ottenute mediante l'analisi ad elementi finiti sono stata sottoposte ad un'analisi di Fourier per renderle confrontabili con i risultati ottenuti con il metodo pertur-

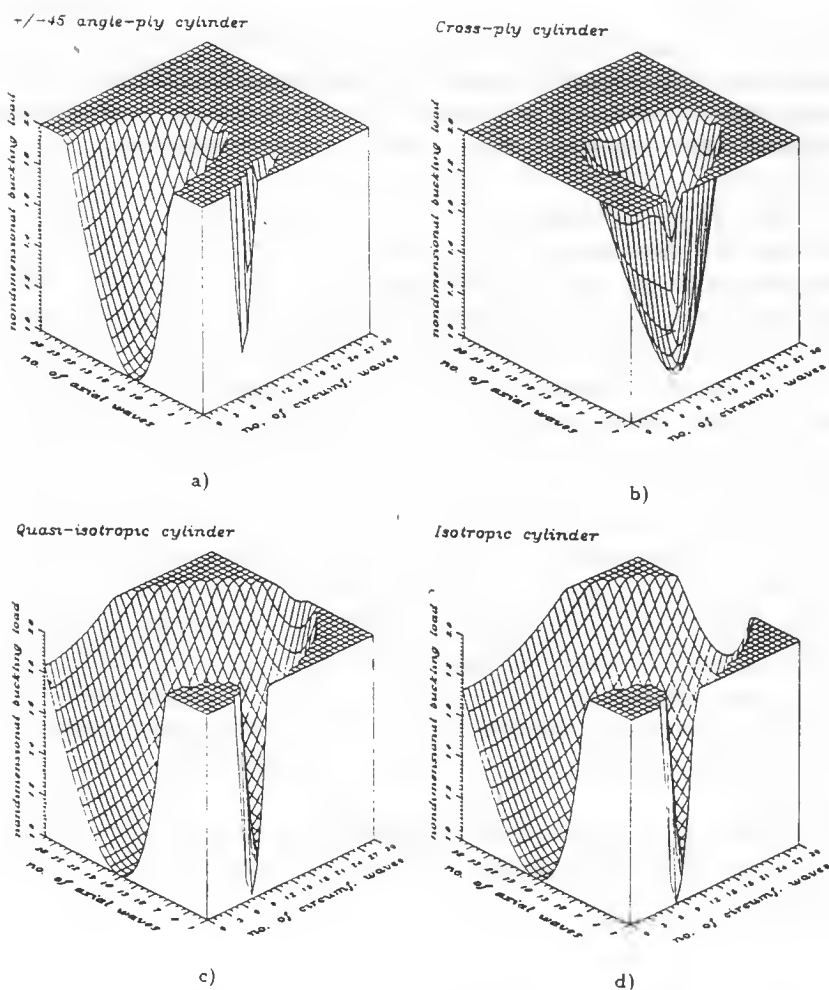


Figura 11 - Carichi critici adimensionali per un cilindro isotropo ed in composito in funzione del numero delle onde circolari e longitudinali.

bativo precedentemente descritto. In fig. 12a-d sono riportati i coefficienti di Fourier ottenuti per le quattro diverse laminazioni esaminate (cross-ply, angle-ply, quasi-isotropo ed isotropo). È immediato notare che sono stati ritrovati gli stessi modi critici ottenuti con il modello teorico. Il modo dominante, ovvero corrispondente al coefficiente di Fourier più alto, per il cilindro cross-ply comprende 6 semionde assiali e 13 onde circolari mentre il modo dominante per il cilindro angle-ply è assialsimmetrico con 13 semionde longitudinali. Nelle fig. 12c-d

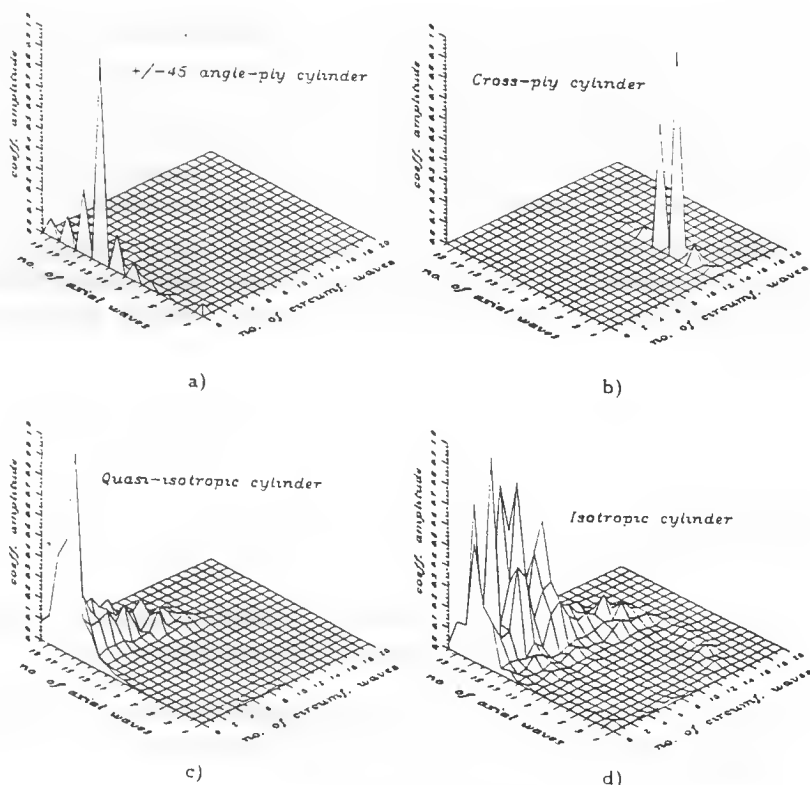


Figura 12 - Analisi di Fourier dei modi critici ottenuti mediante modello ad elementi finiti.

sono riportati invece i risultati ottenuti per un cilindro quasi-isotropo ed uno isotropo, supposto di acciaio. È evidente che la distribuzione dei modi attivi per il cilindro quasi-isotropo è decisamente più estesa dei casi di Fig. 12a, b e molto simile a quella relativa al cilindro isotropo.

La ricerca comprenderà l'analisi di singoli modi di imperfezione e la loro combinazione sul carico limite. La scelta dei modi di imperfezione da esaminare sarà dettata dai risultati dell'analisi statistica precedentemente descritta sviluppata sulla serie di cilindri a disposizione.

6.0. Conclusioni

È stato presentato un progetto di ricerca per lo studio dell'instabilità di gusci in materiale composito che prevede lo sviluppo parallelo di un'in-

dagine numerica, teorica e sperimentale. Sono stati presentati i risultati sperimentali per cilindri cross-ply compressi assialmente. È stata prestata particolare attenzione alla registrazione ed all'esame statistico delle imperfezioni geometriche di due serie di provini al fine di individuarne caratteristiche comuni ed esaminarne l'effetto sul carico critico. È stata inoltre presentata la procedura per la soluzione del problema pre e post-critico di cilindri in composito mediante tecniche perturbative. I risultati ottenuti sono stati confrontati con quelli ottenuti numericamente per mezzo di un codice ad elementi finiti. La ricerca comprenderà l'esame di altri casi di carico quali la torsione e la combinazione di torsione con azione assiale.

BIBLIOGRAFIA

- [1] GIAVOTTO V., «Buckling of composite cylindrical shells in compression and torsion» ICAS-90-4.3.2, 17th ICAS Congress, Stockholm, September 1990.
- [2] ARBOCZ J., «Shell stability analysis: theory and practice» in *Collapse: the Buckling of Structures in Theory and Practice*, J.M.T. Thompson and G.W. Hunt (eds), Cambridge University Press, 1983, pp. 43-74.
- [3] ARBOCZ J., BABCOCK C.D., «Prediction of buckling loads based on experimentally measured imperfections» in *Buckling of Structures*, B. Budiansky (ed), Springer Verlag, 1976, pp. 291-311.
- [4] CHRYSSANTHOPOULOS M.K., BAKER M.J. and DOWLING P.J., «Statistical analysis of imperfections in stiffened cylinders», *J. of Struct. Eng.*, ASCE, 117 (7), 1991, pp. 1979-1997.
- [5] CHRYSSANTHOPOULOS M.K., BAKER M.J. and DOWLING P.J., «Imperfection modelling for buckling analysis of stiffened cylinders», *J. of Struct. Eng.*, ASCE, 117 (7), 1991, pp. 1998-2017.
- [6] ELISHAKOFF I., ARBOCZ J., «Reliability of axially compressed cylindrical shells with random axisymmetric imperfections», *J. Appl. Mech.*, Vol. 52, 1985, pp. 122-128.
- [7] CIAVARELLA M., «Modelli probabilistici di imperfezioni geometriche per lo studio dell'instabilità di gusci cilindrici in materiale composito», Diploma Thesis, Dept. of Aerospace Engineering, Politecnico di Milano, 1991.

- [8] CHRYSSANTHOPOULOS M., GIAVOTTO V., POGGI C., «Statistical imperfection models for buckling analysis of composite shells» in *Buckling of Shells Structures, on land, in the sea and in the air*, J.F. Jullien ed, Elsevier, 1991, pp. 43-52.
- [9] GIAVOTTO V., POGGI C., DOWLING P.J., CHRYSSANTHOPOULOS M., «Buckling Behaviour of Composite Shells under Combined Loading» in *Buckling of Shells Structures, on land, in the sea and in the air*, J.F. Jullien ed, Elsevier, 1991, pp. 53-60.
- [10] SIMITSIS G.J., SHAW D., SHEINMAN I., «Imperfection Sensitivity of Laminated Cylindrical Shells in Torsion and Axial Compression», *Composite Struct.* No. 4, 1985, pp. 335-360.
- [11] ARBOCZ J., HOL J.M.A.M., «Koiter's Stability theory in a computer aided engineering environment», *Int. J. Solids Structures*, Vol. 26, No. 9/10, 1990, pp. 945-973.
- [12] POGGI C., TALIERCIO A., CAPSONI A., «Fiber Orientation Effects on the Buckling Behaviour of Imperfect Composite Cylinders» in *«Buckling of Shells Structures, on land, in the sea and in the air»*, J.F. Jullien ed, Elsevier, 1991, pp. 114-123.

Appendice

Definizione degli operatori differenziali ed algebrici

$$L_1(.) = A_{22}^*(.)_{,xxxx} + 2 \left(A_{12}^* + \frac{1}{2} A_{66}^* \right) (.)_{,xxyy} + A_{11}^*(.)_{,yyyy}$$

$$L_2(.) = B_{21}^*(.)_{,xxxx} + (B_{11}^* + B_{22}^*) (.)_{,xxyy} + B_{12}^*(.)_{,yyyy} + \frac{1}{R} (.)_{,xx}$$

$$L_3(.) = D_{11}^*(.)_{,xxxx} + 2 (D_{12}^* + 2 D_{66}^*) (.)_{,xxyy} + D_{22}^*(.)_{,yyyy}$$

$$N_4(H, K) = H_{,xx} K_{,yy} - 2 H_{,xy} K_{,xy} + H_{,yy} K_{,xx}$$

$$\hat{A}_{mn} = A_{22}^* \left(\frac{m\pi}{L} \right)^4 + 2 \left(A_{12}^* + \frac{1}{2} A_{66}^* \right) \left(\frac{mn\pi}{LR} \right)^2 + A_{11}^* \left(\frac{n}{R} \right)^4$$

$$\hat{B}_{mn} = B_{11}^* \left(\frac{m\pi}{L} \right)^4 + (B_{11}^* + B_{22}^*) \left(\frac{mn\pi}{LR} \right)^2 + B_{12}^* \left(\frac{n}{R} \right)^4$$

$$\hat{D}_{mn} = D_{11}^* \left(\frac{m\pi}{L} \right)^4 + 2 (D_{12}^* + 2 D_{66}^*) \left(\frac{mn\pi}{LR} \right)^2 + D_{22}^* \left(\frac{n}{R} \right)^4$$

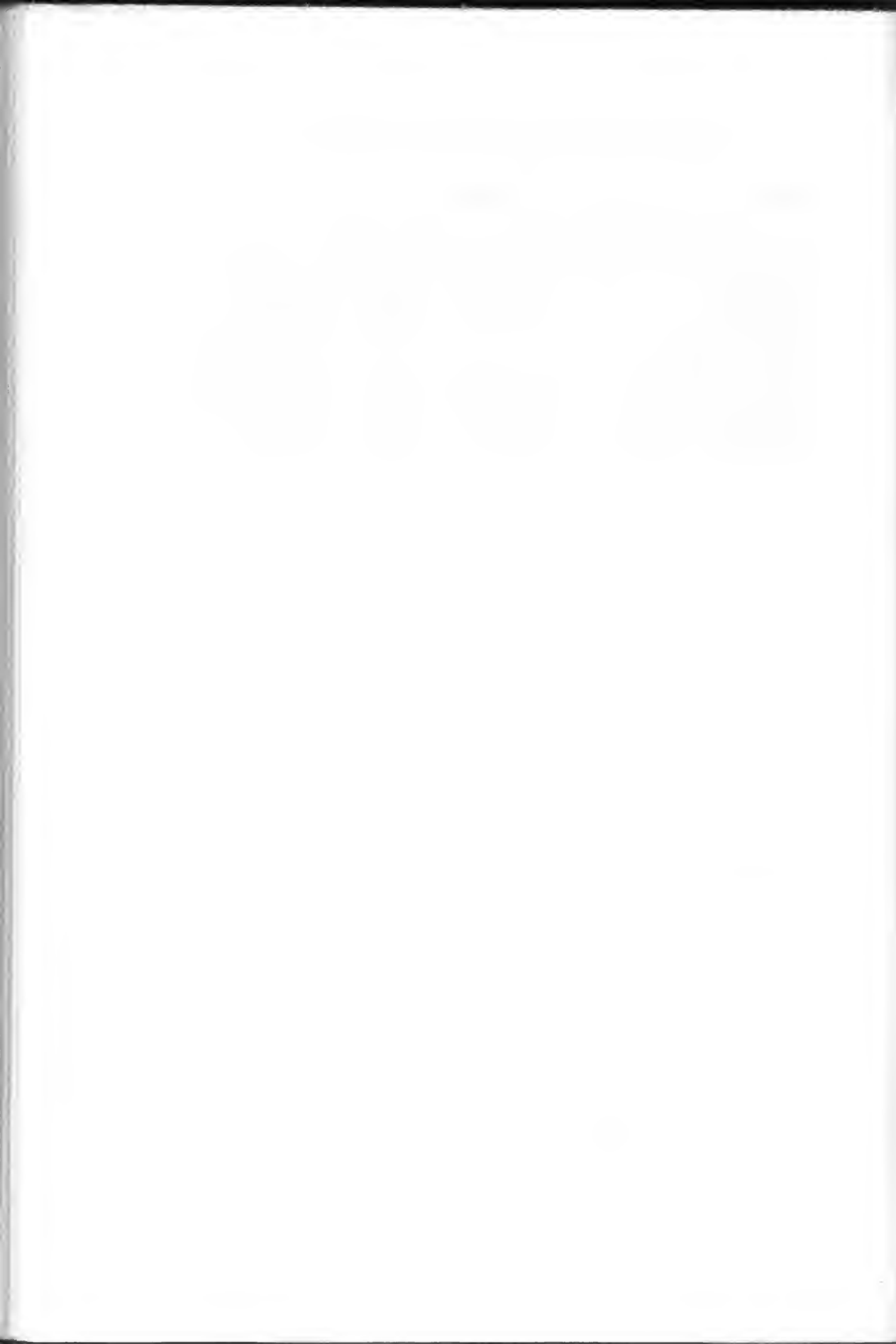
I termini asteriscati indicano gli elementi della matrice di rigidezza del laminato espressa in forma semiinversa.

Tabella 1 - Carichi ultimi sperimentali e carico critico teorico per cilindri perfetti.

Cilindri serie A (cross-ply)		
Cil. no.	Ris. sper.	Sper./Teor.
3	27.14	0.83
5	27.81	0.75
6	28.65	0.77
10	27.77	0.75

Tabella 2 - Carichi critici per cilindri perfetti

Carichi critici per cilindri perfetti (KN)			
Laminazione	El. fin.	eq. (9)	Modo crit.
$(0^\circ/90^\circ)_s$	38.33	37.07	ASIM
$(45^\circ/-45^\circ)_s$	37.03	37.03	ASSIALSIM
$(45^\circ/0^\circ/90^\circ/45^\circ)_s$	70.35	69.93	ASSIALSIM
$(45^\circ/-45^\circ/0^\circ/90^\circ)_s$	66.02	65.71	ASSIALSIM
$(0^\circ/90^\circ/45^\circ/-45^\circ)_s$	70.30	65.75	ASIM



Fluidodinamica della cochlea

Claudio CANCELLI*

1. Introduzione

Non è sorprendente che sia stato Helmholtz il primo a proporre una teoria dell'udito. Riprendendo un'intuizione di Ohm che aveva attribuito all'orecchio il potere di risolvere il segnale acustico nelle sue componenti spettrali, Helmholtz suppose che questa funzione fosse svolta da una struttura interna della *cochlea*, la membrana basilare [1, 1863].

Egli sapeva, o immaginò, che la membrana è costituita essenzialmente da fibre trasversali la cui lunghezza varia in modo ordinato lungo l'asse a spirale della *cochlea*, con le fibre più brevi in prossimità della base e quelle più lunghe all'estremità opposta. Suppose pertanto che le fibre costituissero un insieme di corde vibranti, una batteria dei ben noti risonatori, ciascuno dei quali oscillasse a una frequenza caratteristica diversa.

Nacque così la «teoria della discriminazione spaziale della frequenza», come in seguito è stata comunemente chiamata. Al variare del loro contenuto di frequenza i segnali acustici pongono in vibrazione fibre diverse della membrana basilare e queste a loro volta eccitano terminazioni nervose distinte. L'individuazione soggettiva dell'altezza del suono è legata alla diversità delle terminazioni nervose che vengono eccitate.

Questa interpretazione del ruolo della *cochlea* ha dominato per 60 anni la teoria dell'udito; forse, si potrebbe dire semplicemente che essa è stata in questo periodo la teoria dell'udito.

Ma il carattere della teoria appariva singolarmente in contrasto con la natura dell'oggetto a cui essa si riferiva: l'ipotesi di Helmholtz era semplice e chiara; l'organizzazione della *cochlea* appariva complessa e in molti aspetti funzionali indecifrabile.

Le prime osservazioni del moto della membrana basilare e le prime misure delle proprietà meccaniche di questa e di altre componenti della *cochlea*, dovute a Békésy [2] a partire dal 1928, non contribuirono a chia-

* Dipartimento di Ingegneria Aeronautica e Spaziale, Politecnico di Torino.

rire il quadro. Anzi, esse sembrarono affossare la teoria della risonanza senza suggerirne alcuna altrettanto comprensiva.

Questa situazione si è protratta a lungo; i dati di Békésy sono stati di riferimento per un paio di generazioni di ricercatori e la teoria di Helmholtz è stata di conseguenza ricordata fino a metà degli anni '70 come un tentativo intelligente, ma nei fatti insostenibile, di spiegare il potere di discriminazione dell'apparato uditivo.

Poi, nel volgere di pochi anni, più sofisticate tecniche di misura hanno riportato la teoria della risonanza alla ribalta, in una forma nuova e più complessa di quanto fosse stata originariamente pensata, ma coerente nei lineamenti essenziali con le idee di Helmholtz. I dati di fatto si sono piegati alle esigenze di una costruzione teorica che traeva la sua legittimità, almeno inizialmente, più dalla sua coerenza interna che dalla comprensione delle peculiarità dell'oggetto.

Alcuni aspetti della fisiologia cocleare rimangono da chiarire; ma altri sono stati chiariti e tra essi un interessante problema di meccanica dei fluidi che può essere delineato in questi termini. Nella *cochlea* stimolata si propaga un'onda la cui dinamica è determinata dall'accoppiamento tra il moto dei liquidi contenuti nei canali cocleari e una partizione dotata di proprietà elastiche, inerti e dissipative, che li separa. Come è possibile che l'onda trasporti l'energia fino a una particolare posizione della membrana, dipendente dalla frequenza, e colà la concentri in una zona assai ristretta, senza essere riflessa e senza procedere oltre? Poiché questo è quanto una teoria di risoluzione spaziale della frequenza richiede.

Su quale sia la risposta a questa domanda, e su alcuni aspetti del moto dei fluidi nel canale più interno della *cochlea* non ancora del tutto compresi, verte brevemente questa relazione.

2. Alcuni dati

L'orecchio interno è contenuto in un intricato sistema di cavità e canali entro l'osso temporale ed è costituito da due parti comunicanti che svolgono funzioni distinte: l'apparato vestibolare e la *cochlea*.

L'apparato vestibolare porta i recettori di accelerazione che rendono possibile il controllo dell'equilibrio; la *cochlea* trasduce gli spostamenti del segnale acustico in eccitazione di fibre nervose. Globalmente, la cavità cocleare ricorda quella di una chiocciola, giusto il nome: un tubo rastremato dalla base all'apice con l'asse longitudinale avvolto a spirale (fig. 1). Al suo interno è tuttavia diviso, mediante due membrane, in tre

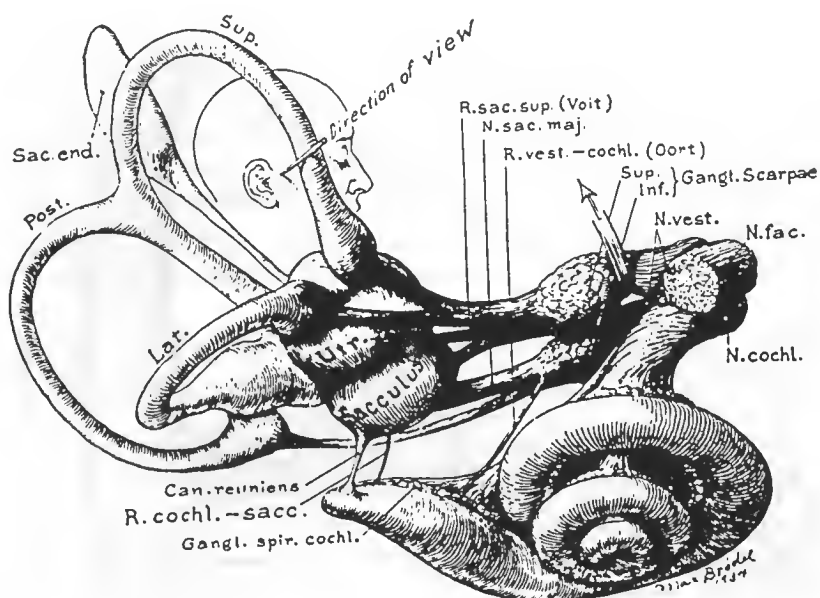


Figura 1 - Orecchio interno.

canali (fig. 2), di cui quello intermedio, la *scala media*, è isolato e termina in un fondo cieco poco prima dell'apice; i due canali esterni, la *scala vestibuli* e la *scala tympani*, sono separati alla base e per gran parte della *cochlea* ma comunicano tra loro in prossimità dell'apice per la scomparsa del canale intermedio. Delle due strutture di divisione, la membrana di Reissner è sottile e flessibile ed ha inerzia e rigidità elastiche del tutto trascurabili; essa sembra avere come unica funzione quella di tenere il liquido interno alla *scala media* separato da quello contenuto nelle *scalae* periferiche — liquido perilinfatico — che ha proprietà elettrochimiche diverse. La partizione che divide la *scala media* dalla *scala tympani* è invece struttura composta, formata dalla membrana basilare a cui si deve la rigidità elastica della partizione, dall'organo di Corti che dalla membrana basilare è sostenuto e che rappresenta il vero organo sensoriale, e dalla membrana tectoria. Quest'ultima è come incernierata alla parete ossea ad un estremo e ricopre l'organo di Corti a cui è connessa mediante le ciglia delle cellule di ricezione — le cellule cigliate — che del Corti costituiscono la struttura essenziale; alle ciglia è affidata la trasduzione della vibrazione meccanica in potenziale di azione nervosa. Nell'insieme, la partizione cocleare ha rigidità elastica e inerzia considerevoli e, molto probabilmente, proprietà dissipative.

Il moto oscillatorio captato dal timpano, e trasmesso dall'orecchio

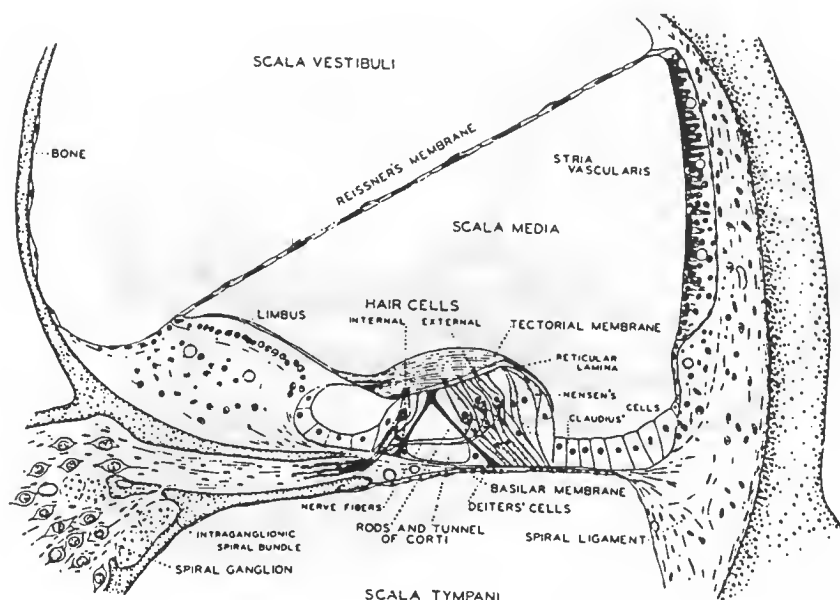


Figura 2 - Sezione retta della *cochlea*.

medio, viene comunicato al liquido della *scala vestibuli* mediante un'apertura membranosa — la finestra ovale — a cui è appoggiata la staffa dell'orecchio medio.

Gli spostamenti del liquido, scarsamente compressibile, che si trova all'interno della *cochlea* sono resi possibili da una seconda apertura — la finestra rotonda — comunicante con la *scala tympani* e sigillata da una membrana flessibile in grado di deformarsi liberamente.

Le prime osservazioni empiriche del moto all'interno della *cochlea* si devono a Békésy (1928), il quale dedicò molti anni allo studio dell'apparato uditivo e pubblicò su questo argomento una notevole serie di lavori, oggi raccolti in un libro [3].

In breve, Békésy scoprì eseguendo esperimenti *post mortem* che nella partizione cocleare si propaga un'onda di spostamenti trasversali con ritardo di fase continuamente crescente dalla base all'apice. Il moto oscillatorio della staffa provoca un flusso pulsante di liquido perilinfatico attraverso la sezione retta della *scala vestibuli*, il quale a sua volta produce l'inflessione laterale della partizione elastica e un flusso in senso opposto nella *scala tympani*, che trova sfogo attraverso la finestra rotonda. L'inflessione della partizione si trasmette alla zona adiacente;

il fenomeno ondulatorio che ne deriva è simile a quello di un'onda di tensione superficiale.

Giudicando a posteriori, non è chiaro perché Békésy abbia ritenuto che la presenza di un'onda progressiva, invece che stazionaria, fosse in sé incompatibile con la teoria della risonanza. L'involuppo delle ampiezze di oscillazione mostrava un massimo, sia pure piatto, che si spostava dalla regione apicale a quella basale (fig. 3) al variare della frequenza, il che non era in radicale disaccordo con il pensiero di Helmholtz.

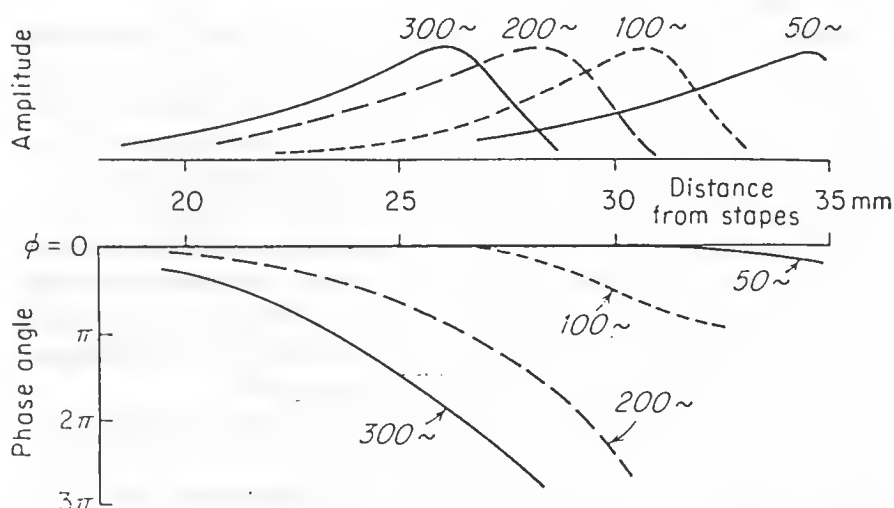


Figura 3 - Ampiezza e fase degli spostamenti trasversali della membrana basilare secondo Békésy (1928); i numeri che caratterizzano le diverse curve rappresentano la frequenza del segnale acustico, misurata in Hz.

I dati quantitativi, invece, erano effettivamente tali da non lasciare credibilità alla teoria della risonanza. Per prima cosa, la membrana basilare appariva esser tutt'altro che una membrana in senso proprio; essa si presentava come una piastra dotata di rigidità flessionale, isotropa nonostante la disposizione ordinata delle fibre; l'ipotesi che essi si comportasse dinamicamente come un insieme di oscillatori indipendenti veniva pertanto destituita di fondamento. Inoltre, sebbene la sua larghezza trasversale risultasse variabile di circa 10 volte dalla base, ove era più stretta, all'apice, la sua rigidità elastica volumetrica per unità di lunghezza assiale variava solo nel rapporto di circa 10^2 . Questo dato sembrava conclusivo; poiché la massa per unità di lunghezza della partizione è quasi costante, la frequenza propria degli eventuali risonatori non po-

teva variare più di 10 volte, quindi in un rapporto tra gli estremi più piccolo, per almeno un ordine di grandezza, di quello tra gli estremi del campo uditivo. Tra la percezione dell'altezza del suono e la localizzazione del risonatore, ammesso che avesse ancora senso parlarne, non poteva sussistere alcuna corrispondenza. Infine, anche la viscosità dei liquidi endococleari e in particolar modo di quello interno alla *scala media*, che Békésy descriveva come gelatinoso, risultava troppo elevata perché la curva di risposta in frequenza degli oscillatori potesse avere un elevato grado di risoluzione.

Békésy si rese conto che i suoi dati sul moto oscillatorio della partizione mal si accordavano con l'elevato potere di discriminazione dell'apparato uditivo, e postulò l'esistenza di un ignoto processo neurofisiologico in grado di rendere più acuta la localizzazione del segnale. La natura di questo secondo filtro è stata materia di indagine e di speculazione per qualche decade; oggi è dubbio che la sua esistenza sia richiesta per accordare la curva di risposta della membrana basilare con la misura diretta dell'eccitazione di singole fibre nervose efferenti dalle cellule cigliate. Delle ricerche di Békésy infatti sono rimasti validi solo gli aspetti qualitativi, mentre i dati quantitativi sono radicalmente cambiati coll'affinarsi delle tecniche di misura [4-13].

Osservati *in vivo* in un punto fisso lungo l'asse della *cochlea* con un livello moderato di stimolazione (Békésy era stato costretto a simulare segnali di intensità estremamente elevata), gli spostamenti della membrana basilare mostrano una curva di risposta in frequenza ben più selettiva di quelle trovate da Békésy. La curva di risposta in un punto fisso (fig. 4) ricorda quella di un filtro passa-basso; è fortemente asimmetrica ed il suo aspetto saliente è dato dalla estrema ripidità di caduta, da 100 a 300 db/ottava, che si ha al di sopra di una frequenza critica il cui valore varia al variare della posizione lungo l'asse. Di converso, un'onda di frequenza determinata si estende soltanto dalla base fino ad una particolare posizione, da cui tuttavia non si riflette se non in misura insignificante; al crescere della frequenza questa posizione si avvicina alla regione basale; il carattere di onda progressiva è confermato dalle misure del ritardo di fase. Per quanto riguarda la rigidità della membrana basilare, essa risulta compresa in un campo più vasto di quello trovato da Békésy, con un rapporto tra gli estremi dell'ordine di 10^3 - 10^4 ; la membrana stessa mostra *in vivo* una forte ortotropia che tende a scomparire rapidamente *post mortem* [14], con una rigidità flessionale longitudinale assai più piccola di quella trasversale. Infine la viscosità dei liquidi endococleari non è molto diversa da quella dell'acqua. Questi sono

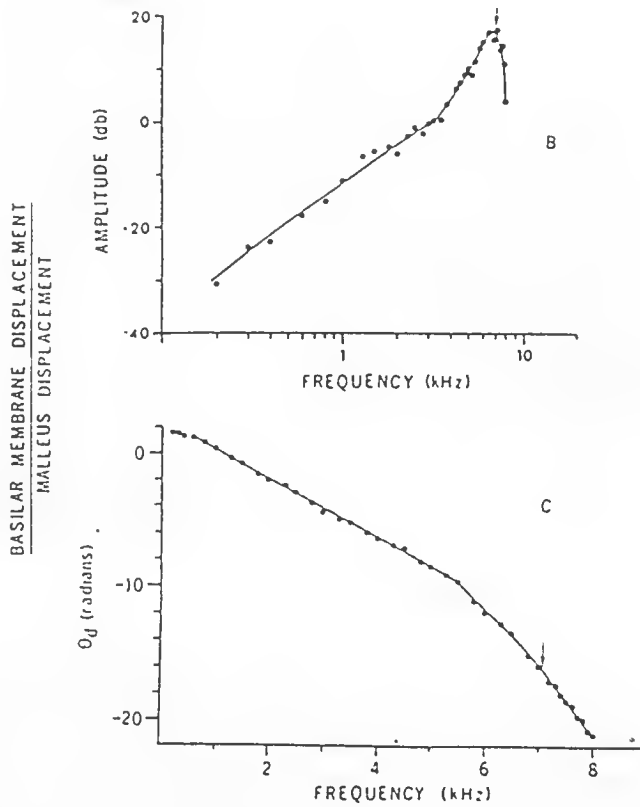


Figura 4 - Ampiezza e fase degli spostamenti trasversali della membrana basilare, misurati in una posizione fissa lungo l'asse della *cochlea*, secondo Rhode (1971).

gli elementi, ormai certi, di cui deve tener conto qualsiasi tentativo di interpretazione teorica del moto.

3. Aspetto matematico del problema

Per quanto possa sembrare strano, né la forma a spirale dell'asse della *cochlea*, né la configurazione più o meno circolare della sua sezione retta hanno grande importanza nel determinare la dinamica dell'onda che si propaga al suo interno. Già Békésy aveva intuito questo fatto [3, p. 407]; studi successivi lo hanno confermato [15, 16, 17, 18].

Un modello schematico come quello illustrato in fig. 5, ad esse retti-

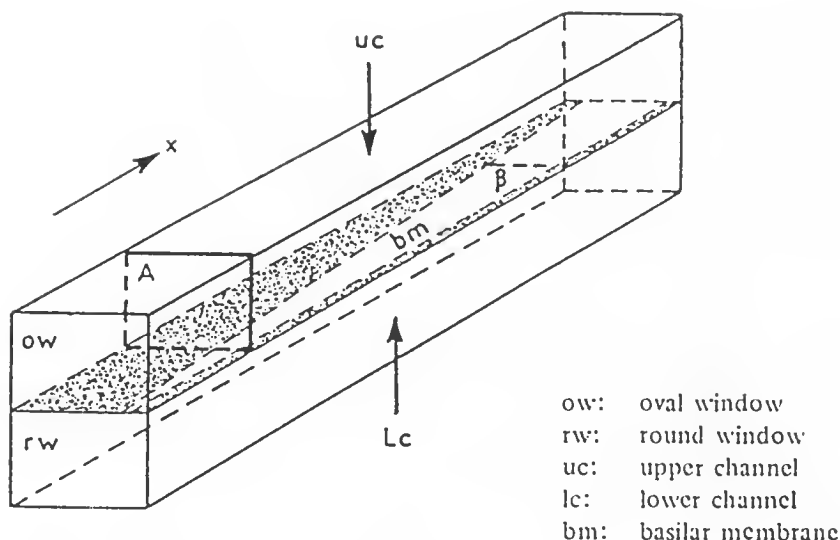


Figura 5 - Modello semplificato di *cochlea*.

lineo e con la sezione reale delle *scala*e sostituita da due sezioni rettangolari di superficie equivalente, conserva gli elementi essenziali del fenomeno di propagazione, purché alla membrana siano attribuite proprietà inerti e dissipative, oltre a quelle elastiche. Il moto del liquido e della partizione elastica è indotto dall'oscillazione della finestra ovale, pilotata dalle staffe.

Il fatto di ignorare la membrana di Reissner e di considerare come identiche le due *scala*e periferiche rende nel modello perfetta una simmetria che nella *cochlea* reale è rispettata solo in modo approssimato. Nel modello si hanno in punti simmetrici rispetto alla partizione spostamenti di fluido, normali alla partizione, identici e spostamenti assiali di uguale grandezza ma tra loro in opposizione di fase. La presenza della *scala tympani* è, da un punto di vista meccanico, del tutto equivalente ad un raddoppio della massa generalizzata del liquido della *scala vestibuli*, qualunque espressione essa abbia. Dunque, anche la complicazione derivante dalla presenza di una partizione immersa nel liquido è solo apparente; il problema, ridotto ai suoi lineamenti essenziali, non è dissimile da un problema classico della idrodinamica: quello di un'onda entro un fluido incompressibile, la cui velocità di propagazione è determinata dall'accoppiamento del fluido con una membrana superficiale. In termini matematici, questo significa che occorre trovare una soluzione

dell'equazione di Laplace:

$$1) \quad \nabla^2 u = 0$$

che soddisfi un insieme di condizioni di contorno, le quali esprimono l'annullarsi della componente della velocità normale alle parti rigide, il moto imposto dalle staffe alla finestra ovale e, infine, l'accoppiamento dinamico tra fluido e membrana. È quest'ultima condizione quella che rende difficile risolvere il problema. Se la membrana viene trattata come una piastra viscoelastica ortotropa, come quasi certamente essa è, la condizione di accoppiamento è espressa da un'equazione differenziale alle derivate parziali, del quarto ordine negli spostamenti della piastra, che non offre speranza di soluzione se non numerica. Ma anche se il problema di quale sia la corretta deformazione della piastra viene aggirato, assumendone a priori il modo di inflessione così che il suo moto venga rappresentato da un'unica coordinata generalizzata, la condizione di accoppiamento rimane severa: una condizione di tipo misto,

$$2) \quad M_p \frac{du_3}{dt} + S \int u_3 dt = \zeta \int_0^b p \zeta dx_2$$

che lega tra loro la componente normale u_3 della velocità del fluido a contatto con la partizione, la massa M_p e la rigidezza generalizzata S di questa, e la pressione p del liquido; pressione e velocità sono tra loro legate dall'equazione di Eulero linearizzata; $\zeta(x_2)$ rappresenta il modo di inflessione trasversale della membrana, soggetto alla condizione normalizzatrice:

$$\int_0^b \zeta(x_2) dx_2 = 1.$$

La (2) non permette una soluzione analitica diretta del problema; e rende difficile anche adottare una delle consuete approssimazioni, che sia valida in tutto il modello, a causa della forte variazione di S lungo l'asse della *cochlea*. Infatti, venendo ai termini propri di un fenomeno di propagazione, occorre notare che la velocità di fase dell'onda risulta grandemente variabile, tra un massimo alla base compreso tra 360 e 50 m/s, ed un minimo tra 5 e 2 m/s all'apice, a seconda delle misure. Ne consegue che per frequenze medie all'interno del campo uditivo la lun-

ghezza d'onda è molto più grande della dimensione lineare tipica della sezione di una scala in prossimità della base (circa 0.7 mm, negli esseri umani), mentre diviene decisamente più piccola di questa in prossimità dell'apice. Dunque, né la consueta approssimazione quasi unidimensionale delle onde in acqua bassa, né quella delle onde in acqua alta è soddisfacente nell'intero dominio e il passaggio dall'una all'altra condizione può avere, e in effetti ha, rilevanza.

Nel tentativo di trovare un accordo soddisfacente con i dati sperimentali, che a loro volta sono andati negli anni variando, quasi tutto il patrimonio di metodi sviluppato per lo studio della propagazione di onde in fluidi incompressibili è stato impiegato per questo caso specifico.

Non possiamo qui, per ragioni di tempo, darne conto; una descrizione accurata di questi metodi può essere trovata nella tesi di dottorato di Viergever [15].

Convien tuttavia ricordare che nella ricerca di metodi matematici che permettano di evitare nello stesso tempo semplificazioni grossolane e un indecifrabile groviglio di calcoli numerici, si è raggiunto infine un compromesso con l'adozione di una tecnica asintotica di alta frequenza, a cui vengono dati nomi diversi (WKB, Liouville - Green), ma che comunque richiede per la sua applicabilità l'ipotesi che la variazione percentuale della lunghezza d'onda λ , che ha luogo lungo l'asse della *cochlea* in una distanza che è dell'ordine di λ stessa, sia piccola rispetto all'unità. Questo presuppone in definitiva che le caratteristiche meccaniche della partizione varino relativamente poco nello spazio di una lunghezza d'onda. Per la maggior parte delle frequenze acustiche questa condizione risulta vera in tutta la *cochlea*, o quasi. Altro fatto notevole è che, indipendentemente dalla sofisticazione formale dei modelli adottati, da essi si possono ottenere risultati molto diversi a seconda delle proprietà meccaniche che si attribuiscono alla partizione; anche modelli assai poco realistici per quanto riguarda il moto del fluido, come quelli che lo assumono a priori essenzialmente longitudinale per tutta l'estensione della *cochlea*, possono dare per l'onda di spostamenti trasversali della membrana risultati assai vicini a quelli misurati *in vivo*.

Quale sia la caratteristica fisica che accomuna i modelli di successo, per così dire, è stato spiegato da Lighthill in un lavoro del 1981 [16]. Lighthill spostò l'attenzione dalle tecniche matematiche al loro significato fisico e propose come criterio unificante di interpretazione l'analisi delle condizioni che presiedono al flusso di energia meccanica lungo la *cochlea*. L'interpretazione di Lighthill è, come vedremo, una riedizione della teoria della risonanza e rappresenta una soddisfacente risposta ai problemi a cui abbiamo accennato nell'introdurre questa nota.

4. La teoria dello strato critico di assorbimento

Lighthill osservò che le condizioni che rendono accettabile il metodo di Liouville - Green assicurano anche che il flusso di energia propagantesi lungo la *cochlea* non si rifletta; in realtà tra i due fatti vi è stretta connessione. Il flusso di energia può allora essere espresso dal prodotto $V_G E$, ove E è l'energia meccanica del moto di oscillazione per unità di lunghezza della *cochlea*, e V_G è la velocità di gruppo dell'onda. Il comportamento della *cochlea* è caratterizzato dal fatto che, per un'onda di frequenza determinata ω , la velocità di propagazione V_G si annulla in una posizione lungo l'asse, che varia al variare della frequenza; crescendo ω , questo punto si sposta verso la base. In un sistema poco smorzato questo fatto comporta un enorme accumulo di energia nello spazio che precede il punto di arresto, dal momento che l'energia dell'onda E

tende a salire come $\frac{1}{V_G}$, finché i fenomeni dissipativi non vengono pre-

potentemente chiamati in causa dal crescere, altrimenti illimitato, dell'ampiezza di oscillazione.

È qui evidente l'analogia con altri fenomeni tipici della propagazione d'onda; in particolare, lo stesso processo è presente nella propagazione di onde interne nell'atmosfera ove la regione dissipativa è indicata con il termine di «strato critico di assorbimento» [20]. L'analisi del processo offre interessanti motivi di riflessione. La velocità di gruppo è determinata dalla relazione di dispersione $\omega(k)$, ove k sta ad indicare il numero d'onda. È il carattere graduale della variazione delle proprietà meccaniche della membrana ciò che permette di definire la relazione di dispersione come una proprietà locale della stessa, continuamente variabile lungo l'asse della *cochlea*. Formalmente, se $H(t, x_1)$ è una qualsiasi misura dello spostamento della membrana, l'energia cinetica T e quella potenziale elastica U del sistema oscillante, per quanto sia complesso, possono scriversi come:

$$T = \frac{1}{2} M \left(\frac{dH}{dt} \right)^2; \quad U = \frac{1}{2} S H^2$$

ove M e S sono la massa e la rigidezza generalizzata del sistema. Ugualizzando la massima variazione di T alla massima variazione di U , si ottiene la relazione di dispersione:

$$3) \quad \omega = \sqrt{\frac{S}{M}}$$

e la velocità di gruppo è data da:

$$V_G = \frac{d\omega}{dk}.$$

La difficoltà, ovviamente, è nell'esprimere S ed M come funzioni di k . Si può, in breve, porre a fuoco quali siano le condizioni essenziali richieste dalla teoria dello «strato critico di assorbimento», dando uno sguardo al diagramma di fig. 6 ove sono riportate alcune curve di dispersione che mostrano l'andamento richiesto da questa teoria: tali da dare

$$\left(\frac{d\omega}{dk}\right) \rightarrow 0, \quad \text{per } k \rightarrow \infty.$$

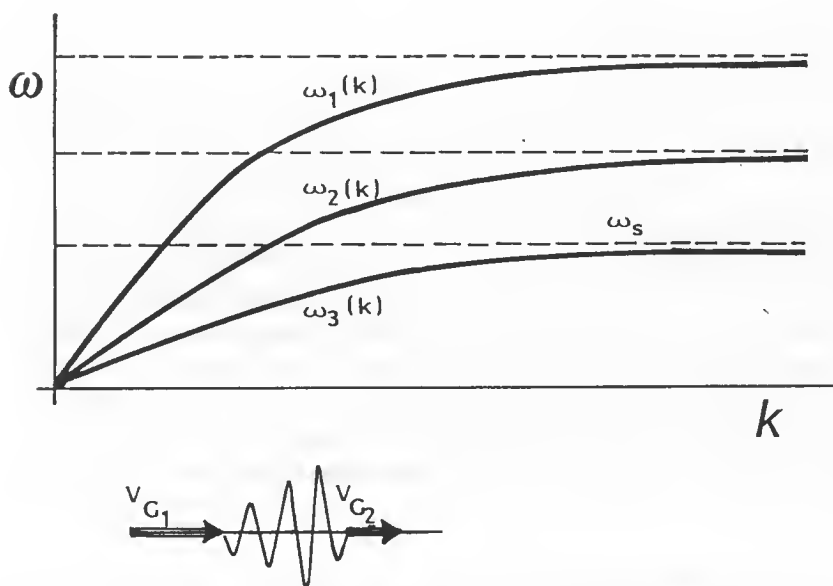


Figura 6 - Curve di dispersione $\omega_1(k)$, $\omega_2(k)$, $\omega_3(k)$, in tre diverse posizioni x_1 , x_2 , x_3 lungo l'asse della *cochlea* ($x_3 > x_2 > x_1$). Un segnale acustico di frequenza angolare ω_s si arresta nella posizione x_3 , ove la velocità di gruppo si annulla; il segnale ha una velocità di gruppo più alta nei punti più vicini alla base, che in figura sono rappresentati da una generica coordinata x_1 , di quanto non abbia nei punti spostati verso l'apice, rappresentati dalla coordinata x_2 .

È facile mostrare come S e M , se dipendono da k , siano l'una funzione crescente, l'altra decrescente di questa grandezza; quindi, l'andamento di fig. 6 ha una sola spiegazione sensata: sia M sia S devono divenire indipendenti dal numero d'onda, al crescere di questo. I modelli matematici che incorporano questa caratteristica sono in grado di dare una descrizione qualitativamente accettabile degli spostamenti della membrana, per quanto grossolani possano essere sotto altri aspetti. La massa generalizzata M è dovuta sia alla partizione, sia al liquido; si può scrivere: $M = M_p + M_L$, separando i due contributi; ma mentre il primo, quello dovuto alla partizione, è indipendente da k , il secondo non lo è ed è quindi necessario che si annulli nel punto di arresto del flusso di energia. Nei modelli quasiunidimensionali, questo avviene perché la compo-

nente assiale della velocità del liquido decresce come $\frac{1}{k^2}$, per ragioni

di continuità, mentre la componente trasversale viene semplicemente ignorata. Nella realtà, e nei più sofisticati modelli a due e tre dimensioni, è il volume di fluido interessato dall'onda che si riduce progressivamente. Nella regione basale un'onda di frequenza acustica coinvolge nel movimento il liquido dell'intera sezione delle *scala*; ma, propagandosi verso l'apice, essa riduce progressivamente il suo raggio di influenza e una di-

stanza dalla membrana sempre più piccola, dell'ordine di $\frac{1}{k}$, finché

la massa del liquido interessato al moto non sia divenuta una percentuale trascurabile della massa della partizione. Qui si ha arresto, senza riflessione, del flusso di energia; e qui è dove la frequenza dell'onda viene a coincidere con la frequenza propria della partizione elastica:

$$\omega_p \sim \sqrt{\frac{S}{M_p}}.$$

La ventura del moto del liquido è, in ultima analisi, quella di svanire nella zona critica, ove si ha il rapido innalzamento e l'ancor più rapida estinzione del moto oscillatorio della membrana. In quanto al massimo dell'oscillazione esso dovrà trovarsi lievemente spostato in senso basale rispetto al punto di risonanza che rappresenta più propriamente il punto estremo di penetrazione dell'onda. Ma dell'originale impostazione di Helmholtz sono rimasti i dati strutturali: la corrispondenza uno ad uno delle frequenze caratteristiche della membrana con la localizzazione dei

massimi e l'idea che la membrana possa essere in qualche modo considerata come una successione ordinata di elementi; poiché non è sostanzialmente diversa da questa l'idea che la membrana, agli effetti del calcolo delle relazioni di dispersione, possa essere considerata nella scala determinata dalla lunghezza d'onda come una successione ordinata di tratti omogenei.

La membrana basilare ha proprietà che appaiono sorprendenti; non essendo tesa essa deve la sua capacità di reazione elastica alla rigidezza flessionale, come una piastra. Ma, se fosse isotropa, questo comporterebbe nell'espressione della rigidezza elastica S dei termini dipendenti anche dalla curvatura longitudinale, quindi da k , che diverrebbero rapida-

mente dominanti per $\frac{kb}{\pi} > 1$ [21], ove b è la larghezza della membrana,

in questo modo impedendo l'arresto nel punto di risonanza del flusso di energia. Ovviamente, la membrana non può avere una rigidezza flessionale elevata in senso trasversale e nulla in senso longitudinale, come la teoria dello «strato di assorbimento critico» a rigore richiederebbe. Ma deve essere sufficientemente ortotropa da permettere che vi sia un campo di valori di kb ove vengano simultaneamente soddisfatte due diverse condizioni: la massa generalizzata del fluido coinvolto nel moto oscillatorio deve essere trascurabile rispetto a quella della partizione e la rigidezza generalizzata della membrana deve risultare tuttora indipendente da k .

Abbiamo già ricordato che in tempi relativamente recenti Voldrich ha scoperto che la membrana ha, *in vivo*, una forte ortotropia che scompare rapidamente *post mortem*. Il rapido decadimento suggerisce una qualche forma di controllo attivo della membrana; vi sono altri fatti che concorrono a sostegno di questa tesi e, in effetti, la ricerca dell'ultimo decennio si è svolta essenzialmente in questa direzione. Torneremo su questo tema nel concludere.

5. Moto dell'endolinfa e modelli fisici di cochlea

Il moto dell'endolinfa all'interno del *sulcus spiralis* e nello spazio compreso tra la membrana tettoria e l'organo di Corti offre motivi di interesse, se non altro per il fatto che le ciglia delle cellule cigliate interne, la cui risposta in frequenza sembra essere più selettiva di quella delle cellule cigliate esterne [22], fluttuano liberamente nel liquido invece di essere infisse nella membrana tettoria. Il dubbio che gli spostamenti

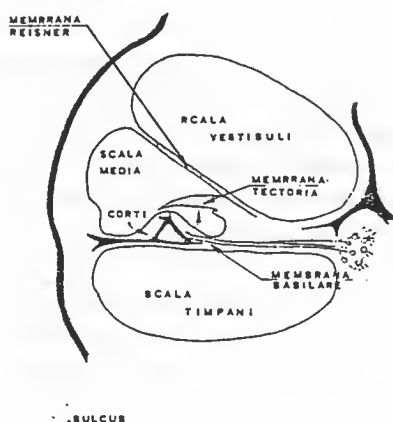
dell'endolinfa all'interno del *sulcus* siano più strettamente localizzati di quelli della membrana sorge spontaneo, sebbene non si veda a priori quale possa esser la spiegazione di un tale fenomeno.

Indipendentemente da questo aspetto il moto delle varie componenti della partizione, tra cui il liquido contenuto nella cavità e nello spazio subettoriale, che è forse comunicante con la *scala media*, è importante in quanto tutti gli elementi concorrono a determinare la massa generalizzata della partizione stessa, che a prima vista risulta troppo piccola per la validità di teorie basate sulla risonanza [23]. Infine, nella *scala media* vi può essere, con maggiore intensità che non nelle *scalae periferiche*, un fenomeno non lineare di origine idrodinamica che è apparso per qualche tempo meritevole di indagine. È noto che un'onda che si propaga in un fluido può generare un flusso continuo, grazie al termine convettivo dell'equazione di Navier-Stokes (cfr. [24], cap. 4.6); questo moto è sostenuto da una sorta di tensore degli sforzi di Reynolds, la cui risultante può essere particolarmente elevata in regioni ove si abbiano intensi gradienti. Strutture vorticosi stazionarie furono osservate da Békésy [3], da Tonndorf [25] e da Helle [26] in modelli fisici di *cochlea*; in particolare Helle osservò un moto continuo di fluido la cui componente trasversale, uscente dal *sulcus*, appariva distribuita in una regione più stretta lungo l'asse del modello di quanto non fosse l'oscillazione della membrana basilare. Poiché il moto continuo è essenzialmente un flusso viscoso, è poco credibile che si distribuisca in uno spazio più limitato di quello occupato dall'oscillazione che lo genera. Tuttavia, la presenza di questo moto poteva essere ritenuta significativa, qualora si fosse postulato che esso potesse avere, grazie al suo carattere continuo, una capacità di saturazione o di inibizione del processo di trasduzione delle cellule cigliate interne. In tal caso, la larga struttura ricircolante del flusso continuo avrebbe inibito l'eccitazione nervosa ovunque fuorché in prossimità del centro del vortice; ed era ragionevole attendersi che il centro si collocasse nella regione del massimo spostamento della membrana, ove i gradienti longitudinali dell'ampiezza del moto oscillatorio invertono in breve spazio il loro segno. Questa interpretazione del tutto fluidodinamica del secondo filtro, ha, a lungo, affascinato chi scrive; il responso dei fatti, come spiegheremo tra breve, non è stato a favore.

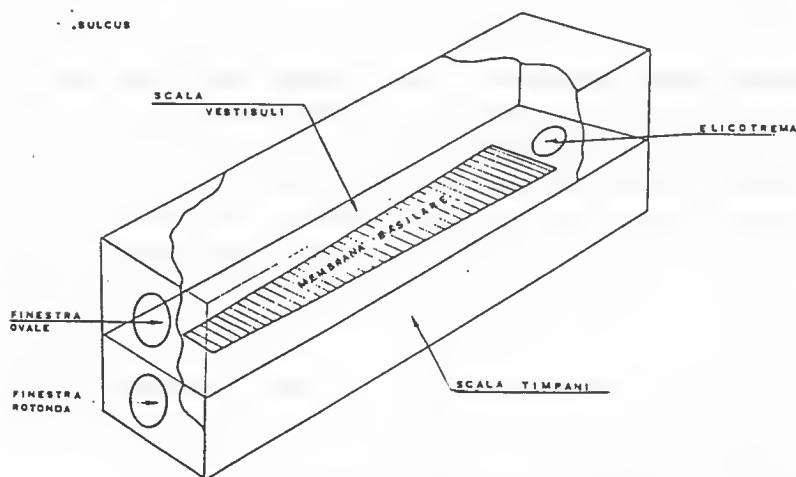
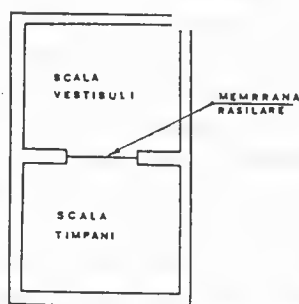
L'indagine su questi temi è stata condotta quasi unicamente in modelli fisici di *cochlea* [26-32], poiché la trattazione teorica appariva troppo complessa e l'osservazione diretta era, e rimane per ora, impossibile. Anche in questo campo Békésy è stato pioniere. Egli costruì a partire dal 1928 alcuni modelli fisici di *cochlea*, e dettò i criteri necessari per assicurare l'identità di forma del moto simulato rispetto a quello reale [2].

Secondo Békésy, il moto della membrana basilare poteva essere riprodotto fedelmente da un modello fisico assai semplice, del tipo di quello illustrato in fig. 7 dove la *scala media* e la struttura reale della partizione vengono semplicemente soppresse. Ma mentre in termini matematici è facilissimo attribuire alla sola membrana le proprietà inerti e dissipative

Sezione retta della coclea



Sezione retta del modello "tipo Bekesy"



Vista spaziale del modello "tipo Békésy"

Figura 7 - Modello fisico di *cochlea* proposto da Békésy.

del sistema reale, in un modello fisico non lo è affatto; i criteri di similitudine proposti da Békésy risultano di fatto incompleti. Un riesame del problema porta rapidamente a concludere che la *scala media* e la partizione vanno riprodotte nel modello [31]; così l'interesse per la dinamica della partizione e la necessità di rispettare i criteri di similitudine si incontrano nel progetto di un modello fisico.

I modelli fisici hanno dato qualche contributo allo studio dei fenomeni oscillatori che avvengono nella *cochlea* — hanno permesso, ad es., di confermare l'applicabilità del metodo L.G. al flusso pienamente tridimensionale — ma non hanno dato sul funzionamento della stessa alcuna idea particolarmente illuminante.

Gli spostamenti di membrana ottenuti sono assai meno selettivi di quelli della *cochlea* reale; anche nel più fortunato dei casi [31] la curva di risposta in frequenza, misurata a punto fisso, mostra una pendenza positiva prima del massimo di $10\text{-}12\text{db/ottava}$, e subito al di là di questo una pendenza negativa di $-40 \div -60\text{ db/ottava}$. Entrambe le pendenze sono in valore assoluto lontane da quelle misurate *in vivo*, e specialmente lo è quella del ramo discendente della curva; in altre parole, nei modelli è risultato difficile riprodurre l'arresto del flusso di energia nel punto di risonanza. È probabile che questo fatto testimoni indirettamente delle singolari proprietà meccaniche della partizione, probabilmente non riprodotte nei modelli. In quanto al moto delle varie componenti della partizione, esso è risultato assai complesso; non è stato tuttavia possibile porre in evidenza alcun aspetto oscillatorio che avesse una distribuzione spaziale più strettamente localizzata dagli spostamenti della membrana [31]. Il moto continuo è realmente presente e in alcuni casi mostra l'andamento qualitativo ipotizzato; ma è, in termini quantitativi, irrilevante finché il segnale viene mantenuto a livelli ragionevoli di intensità.

Simulando nel modello segnali acustici di circa 140 db, si è trovato un rapporto tra la massima velocità del moto continuo e quella del moto oscillatorio dell'ordine di 10^{-2} [31]; poiché il moto continuo è un aspetto non lineare, questo implica che ad un livello di segnale acustico ancora molto alto, diciamo di 80 db, questo rapporto sarebbe all'incirca 10^{-5} . Per quanto uno possa essere ottimista sulla capacità di inibizione di un flusso continuo, non sembra che un'ipotesi di secondo filtro, basata su questo fenomeno, abbia grande fondamento.

6. La membrana basilare come filtro attivo

In realtà, mentre si andavano svolgendo senza grande fortuna le indagini sulle varie ipotesi di secondo filtro, il concetto stesso andava perdendo progressivamente importanza. Le misure più recenti degli spostamenti della partizione, pubblicati nel 1982 da Khanna e Leonard [33] e da Sellik et al. [34] e nel 1986 da Robles et al. [35], vanno oltre i già sorprendenti dati di fig. 4. Si tratta di misure eseguite *in vivo* stimolando l'apparato uditivo con un segnale acustico di bassa intensità, poco al di sopra del valore di soglia; in queste condizioni la curva di risposta in frequenza degli spostamenti di membrana mostra, più o meno, lo stesso andamento dei potenziali elettrocellulari misurati nelle cellule cigliate [36]. Viene da chiedersi se non si possa, almeno per bassa intensità di segnale, dimenticarsi del secondo filtro e assumere semplicemente che l'intensità di eccitazione delle cellule sia una trasformazione lineare della velocità di spostamento della partizione. Con questo torna tuttavia il mistero delle proprietà di quel complesso organo che per brevità abbiamo indicato con il nome di partizione; poiché neppure i modelli matematici che accolgono al loro interno le peculiarità dello strato critico di assorbimento sono in grado di dar conto di una così alta selettività, a meno che non si attribuiscono alla partizione proprietà meccaniche del tutto irrealistiche [37].

La tendenza attuale è quella di considerare la partizione come un organo attivo, o comunque controllato. Vi sono due possibili sviluppi di questa linea di ricerca: uno di essi suppone che la partizione, grazie ad un meccanismo interno non ancora chiarito, trasformi energia chimica in energia meccanica e quindi contribuisca con una sorgente interna al proprio moto. Questa idea era stata proposta in chiave del tutto speculativa da Gold, nel lontano 1948 [38], e quindi dimenticata; è stata ripresa negli anni attorno al 1980 [39, 40, 41, 42], poiché sono stati osservati sperimentalmente fenomeni di emissione acustica da parte della *cochlea* [43], che non sembrano avere origine possibile se non in oscillazioni autoeccitate della partizione. L'altro si basa sull'ipotesi che la partizione possa essere divisa, nel senso della lunghezza, in due fasce adiacenti di proprietà meccaniche diverse, e che l'una di queste fasce che contiene l'organo di Corti e le cellule cigliate possa modificare la propria rigidità elastica in funzione della frequenza del segnale acustico. Questo tentativo di spiegazione, pubblicato recentemente [44, 45], non prevede che la partizione fornisca energia, ma non è meno problematico del primo sotto altri punti di vista; tra l'altro esso tratta le due fasce adiacenti come se fossero meccanicamente disaccoppiate.

È evidente come ipotesi dell'uno o dell'altro tipo aprano un ventaglio di possibilità formali di spiegazione della acutissima curva di risposta della membrana basilare: esse permettono di modificare con funzioni *ad hoc* — resistenza negativa, rigidità variabile con la frequenza etc. — l'impedenza della partizione. Ma è altrettanto ovvio che il criterio discriminante tra le varie ipotesi va cercato nella fondatezza o meno dei processi neurofisiologici invocati. La fluidodinamica della *cochlea* ha perso in queste ricerche il ruolo centrale che ha per lungo tempo avuto; essa appare come un aspetto risolto di un fenomeno più complesso che presenta ancora lati oscuri. Qualche lavoro di natura tipicamente fluidodinamica viene ancora pubblicato [46, 18, 47]; ma sebbene si tratti di lavori tecnicamente interessanti, essi appaiono marginali rispetto alla linea principale di ricerca.

BIBLIOGRAFIA

- [1] HELMHOLTZ H. (1865). *Die Lehre von Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, F. Vieweg und Sohn, Braunschweig.
- [2] BÉKÉSY G. (1928). Zur Theorie des Hörens. *Phys. Z.* 29.
- [3] BÉKÉSY G. (1960). *Experiments in hearing*, McGraw-Hill, N.Y.
- [4] JOHNSTONE B.M. et al. (1970). Mechanics of the guinea pig cochlea. *J. Acoust. Soc. Am.*, 47.
- [5] JOHNSTONE B.M. et al. (1974). Basilar membrane tuning curves in the guinea pig. *J. Acoust. Soc. Am.*, 55.
- [6] RHODE W.S. (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *J. Acoust. Soc. Am.*, 49.
- [7] RHODE W.S. (1973). An investigation of post-mortem cochlear mechanics using the Mössbauer effect. *Basic mechanisms in hearing*, A.R. Møller, Academic, N.Y.
- [8] ROBLES L. et al. (1976). Transient response of the basilar membrane measured in squirrel monkeys using the Mössbauer effect. *J. Acoust. Soc. Am.*, 59.

- [9] WILSON J.P., JOHNSTONE J.R. (1972). Capacitive probe measures of basilar membrane vibration. *Hearing theory 1972*, IPO, Eindhoven.
- [10] WILSON J.P., JOHNSTONE J.R. (1975). Basilar membrane and middle-ear vibration in guinea pig measured by capacitive probe. *J. Acoust. Soc. Am.*, 57.
- [11] KOHLLÖFFEL L.U. (1972 a). A study of basilar membrane vibrations I. Fuzziness-detection: a new method for the analysis of microvibrations with laser light. *Acustica*, 27.
- [12] KOHLLÖFFEL L.U. (1972 b). A study of basilar membrane vibrations II. The vibratory amplitude and phase pattern along the basilar membrane (post-mortem). *Acustica*, 27.
- [14] VOLDRICH L. (1978). Mechanical properties of the basilar membrane. *Acta Otolaryng.*, 86.
- [15] VIERGEVER M.A. (1978). Basilar membrane motion in a spiral-shaped cochlea. *J. Acoust. Soc. Am.*, 64.
- [16] LIGHTHILL J. (1981). Energy flow in the cochlea. *J. Fluid Mechanics*, 106.
- [17] CHEE HOONG LOH (1983). Multiple scale analysis of the spirally coiled cochlea. *J. Acoust. Soc. Am.*, 74.
- [18] STEELE C.R., ZAIS J.G. (1985). Effect of coiling in a cochlear model. *J. Acoust. Soc. Am.*, 77.
- [19] VIERGEVER M.A. (1980). *Mechanics of the inner ear*. Delft University Press.
- [20] BOOKER J.R., BRETHERTON F.P. (1967). The critical layer for internal gravity waves in a shear flow. *J. Fluid Mechanics*, 27.
- [21] STEELE C.R., TABER L.A. (1979). Comparison of WKB and finite difference calculations for a two-dimensional cochlear model. *J. Acoust. Soc. Am.*, 65.
- [22] RUSSELL I.J., SELICK P.M. (1978). Intracellular studies in the mammalian cochlea. *J. Physiol.*, 284.
- [23] ZWISLOCKI J. (1953). Review of recent mathematical theories of cochlear dynamics. *J. Acoust. Soc. Am.*, 25.
- [24] LIGHTHILL J. (1978). *Waves in fluids*. Cambridge University Press. Cambridge.

- [25] TONNDORF J. (1958). Harmonic distortion in cochlear models. *J. Acoust. Soc. Am.*, 30.
- [26] HELLE R. (1974). Enlarged hydromechanical cochlea model with basilar membrane and tectorial membrane. *Facts and models in hearings*, Zwicker and Terhardt, Springer, N.Y.
- [27] TONNDORF J. (1957). Fluid motion in cochlear models. *J. Acoust. Soc. Am.*, 29.
- [28] TONNDORF J. (1958). Beats in cochlear models. *J. Acoust. Soc. Am.*, 31.
- [29] TONNDORF J. (1959). Dimensional analysis of cochlear models. *J. Acoust. Soc. Am.*, 32.
- [30] CANNELL J.K. (1969). *Cochlear models*. Ph. D. thesis, University of Warwick.
- [31] CANCELLI C. et al. (1985). Experimental results in a physical model of the cochlea. *J. Fluid Mechanics*, 153.
- [32] FROMMER G.H. (1977). *Fluid motion in the reticular lamina*. Ph. D. thesis, Stanford University, Palo Alto.
- [33] KHANNA S.M., LEONARD D.G. (1982). Basilar membrane tuning in the cat cochlea. *Science*, 215.
- [34] SELLIK P.M. et al. (1982). Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. *J. Acoust. Soc. Am.*, 72.
- [35] ROBLES L. et al. (1986). Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves and response phases. *J. Acoust. Soc. Am.*, 80.
- [36] SELLIK P.M. et al. (1983). Comparison between the tuning properties of inner hair cells and basilar membrane motion. *Hear. Res.*, 10.
- [37] BOER E. (1983). No sharpening? A challenge for cochlear mechanics. *J. Acoust. Soc. Am.*, 73.
- [38] GOLD T. (1948). Hearing II. The physical basis of the action of the cochlea. *Proc. Roy. Soc. B.*, 210.
- [39] KIM D.O. et al. (1980). An active cochlear model with negative damping in the partition: comparison with Rhode's ante and post-mortem observations. *Psychophysical, physiological and behavioural studies in hearing*, Delft Univ. Press.

- [40] NEELY S.T., KIM D.L. (1983). An active cochlear model showing sharp turning and high sensitivity. *Hearing Res.*, 9.
- [41] NEELY S.T., KIM D.O. (1985). A model for active elements in cochlear biomechanics. *J. Acoust. Soc. Am.*, 79.
- [42] BOER E. (1983). On active and passive cochlear models. Towards a generalized analysis. *J. Acoust. Soc. Am.*, 73.
- [43] KEMPT D.T. (1978). Stimulated acoustic emissions from within the human auditory system. *J. Acoust. Soc. Am.*, 64.
- [44] KOLSTON P.J. (1988). Sharp mechanical tuning in a cochlear model without negative damping. *J. Acoust. Soc. Am.*, 83.
- [45] KOLSTON P.J. et al. (1989). Realistic mechanical tuning in a micromechanical cochlear model. *J. Acoust. Soc. Am.*, 86.
- [46] KELLER J.B., NEU J.C. (1985). Asymptotic analysis of a viscous cochlear model. *J. Acoust. Soc. Am.*, 77.
- [47] FUHRMANN E. et al. (1987). Wave propagation in the cochlea: effects of Reissner's membrane and non-rectangular cross-section. *Acta Mechanica*, 70.

A.S.
5/84